

**Masarykova univerzita
Filozofická fakulta**

Katedra filosofie

Filozofie

Jan Štěpánek

**Tarského definice pojmu pravdy a její
kritika**

Bakalářská diplomová práce

Vedoucí práce: PhDr. BcA. Jiří Raclavský, Ph.D.

2010

Prohlašuji, že jsem diplomovou práci vypracoval samostatně s využitím uvedených pramenů a literatury.

.....

Velmi děkuji PhDr. BcA. Jiřímu Raclavskému, Ph.D. za vstřícnost, cenné rady, podnětné konzultace a trpělivost při vedení mé práce.

Obsah

Úvod	5
Sémantická koncepce pravdy a její důsledky	7
1 Předpoklady adekvátní teorie pravdy	7
1.1 Tarského teoretické východisko	7
1.2 Aplikace termínu „pravdivý“ na věty	9
1.3 Závislost pravdy na jazyku	11
1.4 Kritérium formální správnosti	12
1.5 Kritérium materiální adekvátnosti	14
2 Konvence T	16
2.1 Konvence T a T-schéma	16
2.2 Rozšíření T-schématu	18
2.3 Dva druhy jmen vět	19
3 Tarského dělení jazyků	22
3.1 Jazyky se specifikovanou strukturou, formalizované jazyky	22
3.2 Sémanticky uzavřené jazyky	23
3.3 Status hovorových jazyků	25
3.4 Objektový jazyk a hierarchie metajazyků	26
4 Konstrukce definice pravdy	30
5 Adekvátnost pro řešení paradoxů	33
Kritika Tarského sémantické koncepce pravdy	34
6 Problémy hierarchie metajazyků	34
6.1 Mstivý lhář a nedostatečnost hierarchií metajazyků	34
6.2 Nerozhodnutelnost zařazení k metajazyku	36
7 Záchrana autoreferenčních vět	40
8 Nekonzistence hovorových jazyků	44
9 Aplikovatelnost predikátu „pravdivý“ na věty	47
10 Saul Kripke — proti bivalenci a relativizaci	50
Závěr	54
Použitá a citovaná literatura	56

Úvod

Většina filosofů by souhlasila, že dílo Alfreda Tarského týkající se pravdy bylo nesmírně vlivné. Na tomto místě však shoda obecně končí, neboť ústřední téma Tarského teorie, její aplikovatelnost na filosofické záležitosti a její samotná podstata jsou všechno zevrubně diskutovaná témata.¹

Alfred Tarski patří bezesporu mezi nejvlivnější logiky dvacátého století, a ačkoliv on sám se považoval spíše za matematika a logika než za filosofa,² jeho dílo se stalo velmi inspirativním právě pro filosofy. Jednalo se pak především o jeho práce na poli sémantiky a teorie pravdy, které se dočkaly velkého ohlasu.³ Mým záměrem v této práci je prozkoumat právě tu část Tarského díla, která je zaměřena na teorii pravdy.

Při svých zkoumáních budu vycházet ze dvou nejslavnějších Tarského statí věnovaných problematice pravdy. Tou první je „The Concept of Truth in Formalized Languages“,⁴ (dále jen [CTFL]) která poprvé vyšla roku 1933 v polštině, ale postupem času se dočkala německého a posléze i anglického překladu. Tato práce se však nijak výrazněji nerozšířila v povědomí filosofů a Tarski tento její neúspěch přičítal jejímu příliš technickému charakteru.⁵ Proto se rozhodl svoje výsledky prezentovat ještě jednou, tentokrát ale ve „filosoficky přívětivější“ formě. Výsledkem jeho snažení se stala stať „The Semantic Conception of Truth: and the Foundations of Semantics“,⁶ (dále jen [SCT]) která vyšla poprvé roku 1944. Tarski se v tomto článku omezil na shrnutí nejdůležitějších principů své sémantické koncepce pravdy a vynechal specializované partie věnované konstrukci definice pravdy pro jazyky konečného nebo nekonečného řádu, kterým byla věnována podstatná část [CTFL]. Kromě toho je v [SCT] obsažena polemická část, ve které se Tarski vypořádává s kritikami proti jeho koncepci vznesenými.

¹ „Most philosophers would agree, that Alfred Tarski’s work on truth has been immensely influential. But that is where agreement generally ends, for the point of Tarski’s theory, its applicability to philosophical concerns, and its very nature are all deeply disputed.“ (LYNCH, Michael B. *The Nature of Truth: Classic and Contemporary Perspectives*. MIT Press: Cambridge, 2001. s. 323.)

² TARSKI, Alfred. Sémantická koncepce pravdy a základy sémantiky. In *Logika 20. století: Mezi filosofii a matematikou: výbor textů k moderní logice*. Praha: Filosofia, 2006, s. 173.

³ Srovnej SUPPES, Patrick. Philosophical Implications of Tarski’s Work. *The Journal of Symbolic Logic*. March 1988, vol. 53, no. 1, s. 80–91 a HECK, Richard G. Tarski, Truth, and Semantics. *The Philosophical Review*. October 1997, vol. 106, no. 4, s. 533–534.

⁴ TARSKI, Alfred. The Concept of Truth in Formalized Languages. In *Logic, Semantics, Metamathematics*. Oxford, New York: Oxford University Press, 1956. s. 152–278.

⁵ Viz TARSKI, Alfred. Sémantická koncepce pravdy a základy sémantiky. s. 135.

⁶ TARSKI, Alfred. The Semantic Conception of Truth: and the Foundations of Semantics. *Philosophy and Phenomenological Research*. March 1944, vol. 4, no. 3, s. 341–376. Pro účely této práce používám citace z českého překladu TARSKI, Alfred. Sémantická koncepce pravdy a základy sémantiky. In *Logika 20. století: Mezi filosofii a matematikou: výbor textů k moderní logice*. Praha: Filosofia, 2006. s. 135–176.

Moje práce bude mít po vzoru [SCT] rovněž dvě části — ta první bude věnována výkladu Tarského sémantické koncepce pravdy, druhá pak vybraným kritikám proti této koncepci namířeným. Ve výkladové části bude mým hlavním zdrojem především [SCT], přičemž na některých místech se budu snažit doplnit výklad o informace obsažené pouze v [CTFL]. Druhá a kratší část této práce je věnována některým kritikám vzneseným proti sémantické koncepci pravdy, přičemž se jedná o kritiky, které jsou většinou novějšího data, než kritiky, se kterými se vyrovnává sám Tarski v [SCT].

Sémantická koncepce pravdy a její důsledky

1 Předpoklady adekvátní teorie pravdy

1.1 Tarského teoretické východisko

Termín „pravdivý“, tak jak je používán v každodenní konverzaci v hovorovém jazyku, má nejednoznačný význam. Už od antiky se filosofové, a nejen oni, pokoušejí uspokojivým způsobem vysvětlit, v čem vlastně pravda spočívá a jaký tedy vlastně je význam termínu „pravdivý“. Vzhledem ke skutečnosti, že tento termín je v každodenní konverzaci používán velmi často, mohl by být leckdo sváděn k myšlence, že jeho definování nebude obtížné. Tuto domněnku by snad mohla podporovat i skutečnost, že už jenom schopnost používat v rozhovorech termín „pravdivý“ předpokládá alespoň minimální porozumění tomuto termínu. Sociolog by mohl navrhnout, že průzkum veřejného mínění by mohl napovědět, jak chápe pravdu reprezentativní vzorek populace. Otázkou však je, zdali by takový průzkum vypověděl něco relevantního o povaze pravdy a ulehčil nám tak naše zkoumání. Zdá se, že nikoli, neboť jak Tarski uvádí v polemické části [SCT], „pouze 15% dotázaných lidí souhlasilo s tím, že pojem „pravdivý“ pro ně znamená „souhlasící s realitou“, zatímco 90% respondentů souhlasilo s tím, že věta jako „sněží“ je pravdivá tehdy a jen tehdy, když sněží“.⁷ Avšak i kdyby existovala nějaká možnost spolehnout se na výsledky dosažené touto dotazníkovou metodou, stejně by nebyly s ohledem k záměrům Tarského zkoumání relevantní. V úvodu k [CTFL] Tarski uvádí cíl své práce následovně: „Důkladná analýza významu termínu „pravdivý“ přítomného v každodenním životě zde není zamýšlena. [...] Budu se zabývat výhradně uchopením intencí obsažených v takzvané *klasické* koncepci pravdy.“⁸

Z citovaného je evidentní, že Tarski si mezi různými teoriemi pravdy vybral teorii korespondenční, jejíž tradice vychází z Aristotela (proto o ni Tarski mluví jako o *klasické* koncepci pravdy).⁹ Tím jsou také dány závazky, kterým bude muset jeho koncepce vyhovovat, aby byla v souladu s aristotelským pojetím. Formulaci Aristotelovy koncepce pravdy, ke které Tarski odkazuje v obou svých stěžejních pracích věnovaných pravdě, je

⁷ TARSKI, Alfred. Sémantická koncepce pravdy a základy sémantiky. s. 162.

⁸ „A thorough analysis of the meaning current in everyday life of the term ‘true’ is not intended here. [...] I shall be concerned exclusively with grasping the intentions which are contained in the so-called *classical* conception of truth.“ (TARSKI, Alfred. The Concept of Truth in Formalized Languages. s. 153.)

⁹ Důsledky Tarského rozhodnutí zabývat se již existující teorií pravdy a nikoli vytvářením úplně nové se budu detailněji zabývat v kapitole 1.5.

možné najít ve čtvrté knize *Metafyziky* (konkrétně v kapitole věnované výkladu zákona vyloučeného třetího):

- (1) Omylem jest, jestliže řekneme, že jsoucí není anebo že nejsoucí jest; pravdou jest, řekneme-li, že jsoucí je a nejsoucí není.¹⁰

Třebaže Aristotelovu definici pravdy uvádí Tarski v obou svých hlavních spisech věnovaných pravdě, tak pouze v [SCT] ji srovnává s ostatními, modernějšími formulacemi korespondenční teorie pravdy:

- (2) Pravdivost věty spočívá v jejím souhlasu (či korespondenci) se skutečností.
- (3) Věta je pravdivá, designuje-li existující stav věcí.¹¹

Další moderní formulaci korespondenční teorie pravdy nabízí Tarski ve své dřívější a mnohem techničtější stati [CTFL]:

- (4) [P]ravdivá věta je taková věta, která říká, že stav věcí je takový a takový, a stav věcí je vskutku takový a takový.¹²

Tarski v [SCT] uvádí, že Aristotelova definice je mnohem jasnější a přesnější než novodobé formulace, ale ani ona není bezchybná. Jeho zálibu v Aristotelově klasické koncepci pravdy je možné vysvětlit tradicí, ze které vyšel — jeho učitel, Tadeusz Kotarbiński byl žákem Kazimierze Twardowského a ten byl zase žákem Franze Brentana, který se specializoval, mimo jiné, i na Aristotelovo učení.¹³

Podrobnější analýzu těchto formulací korespondenční teorie nabízí Donald Davidson,¹⁴ který stejně jako Tarski pokládá za nejlepší z uvedených čtyř formulací tu Aristotelovu.¹⁵ Oba se však rozcházejí ohledně hodnocení její obecnosti — Davidson totiž nesdílí Tarského názor, že Aristotelova definice není dostatečně obecná. Aristotelova

¹⁰ *Met.* Γ 1011b26. Používám překlad z ARISTOTELÉS. *Metafyzika*. Přeložil Antonín Kříž. Praha: Nakladatelství Petr Rezek, 2003.

¹¹ Jak věta (2), tak věta (3) pocházejí z TARSKI, Alfred. *Sémantická koncepce pravdy a základy sémantiky*. s. 138.

¹² „[A] true sentence is one which says that state of affairs is so and so, and the state of affairs indeed is so and so.“ (TARSKI, Alfred. *The Concept of Truth in Formalized Languages*. s. 155.)

¹³ Viz MILNE, Peter. *Tarski on Truth and Its Definition* [online]. 1997 [cit. 2010-02-12]. Dostupné z: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.103.712&rep=rep1&type=pdf>.

¹⁴ DAVIDSON, Donald. *The Folly of Trying to Define Truth*. *The Journal of Philosophy*. June 1996, vol. 93, no. 6, s. 263–278.

¹⁵ Nicméně Tarski explicitně uvádí pouze to, že Aristotelova *klasická* formulace (1) předčí moderní formulace (2) a (3). Do [CTFL] byla Aristotelova formulace doplněna až v roce 1956, tedy po vydání [SCT]. Z této skutečnosti a dále z toho, že Tarski nepoužil v Sémantické koncepci pravdy znovu formulaci (4) a že v roce 1956 nepřidal do [CTFL] formulace (2) a (3), je dle mého názoru oprávněné vyvodit, že Tarski považoval formulace (2–4) za víceméně rovnocenné. Soudím, že tvrzení, že Tarski se shoduje s Davidsonem v „glorifikování“ Aristotelovy formulace, je oprávněné.

definice je ostatně Davidsonem vyzdvihována jako velmi zdařilá, přičemž v její prospěch jsou uvedeny tři argumenty. Zaprvé: definice (3) a (4) odkazují ke stavům věcí [state of affairs], které jsou podle Davidsona v sémantice k ničemu.¹⁶ Zadruhé: Aristotelova formulace neobsahuje „podivné mezery“ jako je „takový a takový“ které se vyskytují v Tarského formulaci (4).¹⁷ Třetím argumentem je pak skutečnost, že z Aristotelovy formulace, na rozdíl od ostatních tří, je zřetelné, že pravdivost věty záleží na struktuře této věty. Osobně bych přidal ještě čtvrtý důvod — Aristotelova definice neobsahuje žádné problematické pojmy, jako jsou např. pojmy „designuje“ z definice (3) nebo „koresponduje“ z definice (2).

1.2 Aplikace termínu „pravdivý“ na věty

V této kapitole se zmíním o tom, na jaké entity je podle Tarského aplikovatelný termín „pravdivý“. Vhodné bude začít vysvětlením pojmu pravděnec,¹⁸ což je to, čemu je přisuzována pravdivostní hodnota, respektive to, co může pravdivostní hodnotu nabývat.¹⁹ V každodenním životě a hovorovém jazyku nemá termín „pravdivý“ jasně stanovené vymezení — ostatně např. v angličtině je možné termín „true“ použít i ve slovním spojení „true friend“, které nemá s problematikou pravdy příliš mnoho společného. Logikové a filosofové nám jednoznačnou odpověď na otázku, co lze jednoznačně označit za pravděnec, také neposkytnou — „Mezi kandidáty jsou přesvědčení, propozice, soudy, tvrzení, výroky, teorie, zmínky, ideje, akty myšlení, prohlášení, větné tokeny, větné typy, věty (nespecifikované) a řečové akty.“²⁰

Vzhledem ke skutečnosti, že se Tarski snaží o vybudování teoretické sémantiky, kterážto je součástí logiky, mohl by někdo navrhnout, aby byl termín „pravdivý“ aplikován na propozice. Tuto možnost sice Tarski uvažuje, ale okamžitě ji zavrhuje, jelikož shledává pojem propozice příliš mnohoznačným a jeho význam kolísavým.

Pokud jde o termín „propozice“, je jeho význam neblaze proslulým tématem rozvláčných rozepří mezi mnoha filosofy a logiky a nezdá se, že by byl kdy vysvětlen jasně a jednoznačně.

¹⁶ Tento argument by se snad dal zobecnit tak, že Aristotelova definice nezavádí žádné nové a problematické entity.

¹⁷ Viz DAVIDSON, Donald. *The Folly of Trying to Define Truth*. s. 267.

¹⁸ *Pravděnec* je v českém prostředí již poměrně zažitý překlad anglického termínu *truth-bearer*, se kterým přišel prof. Pavel Materna.

¹⁹ To, díky čemu se pravděnec stává pravdivým, je označováno termínem *truth-maker*, pro který se vžil překlad prof. Materny *pravditel*.

²⁰ „Among the candidates are beliefs, propositions, judgements, assertions, statements, theories, remarks, ideas, acts of thought, utterances, sentence tokens, sentence types, sentences (unspecified), and speech acts.“ (KIRKHAM, Richard L. *Theories of Truth: A Critical Introduction*. Cambridge: MIT Press, 1995. s. 54.)

Z několika důvodů se jeví nejvhodnějším aplikovat termín „pravdivý“ na věty, a my se této praxi přidržíme.²¹

V době, kdy Tarski psal své dvě hlavní práce zabývající se pravdou, nebyl pojem propozice jednoznačně definován a Tarského odpor vůči němu je tak do značné míry oprávněný.

Tarského postoj vůči entitám hrajícím v jeho koncepci roli pravděnců shrnuje Peter Milne následovně: „[P]ro Tarského je pravda vypovídána [...] o větách, nikoliv o obsazích vět, [...] ani o větách společně s jejich obsahy.“²² Pravda je tedy v Tarského koncepci vždy pravdou vět, nikoli pravdou jejich obsahu/významu. Jak však bylo uvedeno v citátu na počátku této kapitoly, vzniká zde otázka, jestli myslel Tarski větné typy, tokeny nebo prostě věty, které již dále nijak nespécifikoval. Odpověď na tuto otázku bohužel není jednoznačná. V [SCT] Tarski exemplifikuje pojem věty tak, že ho bude pojímat „ve smyslu tříd nápisů podobného tvaru“.²³ Toto velmi dobře odpovídá tradičnímu pojetí větného typu tak, jak ho definuje Kirkham.²⁴ Ten však o Tarském dále uvádí, že nejdříve byl sice Tarski zastáncem názoru, že roli pravděnců hrají větné typy, později v životě však změnil názor a začal coby pravděnce ve své koncepci používat větné tokeny.²⁵ Tento „obrat k tokenům“ však nastal až na konci šedesátých let. Jelikož v této práci vycházím z Tarského děl, vydaných před tímto „obratem“, je vhodné pojem věty pro účely této práce chápat jako větný type. Je však také nutné vzít v úvahu, že Tarski se nebránil tomu, aby byla jeho koncepce pravdy uzpůsobena tak, aby v ní mohly hrát roli pravděnců i jiné entity, což explicitně vyjadřuje v [SCT]: „To, že nás zde v první řadě zajímá pojem pravdy pro věty, samozřejmě nevylučuje možnost jeho následovného rozšíření i na jiné druhy předmětů.“²⁶

Problematikou aplikovatelnosti termínu „pravdivý“ na věty se budu dále zabývat ve druhé části této práce, kde představím koncepci Jonathana Harrisona napadající Tarského aplikování termínu „pravdivý“ na věty.

²¹ TARSKI, Alfred. *Sémantická koncepce pravdy a základy sémantiky*. s. 137.

²² „[F]or Tarski truth is predicated, [...] of sentences, not of the contents of sentences nor [...] of sentences jointly with their contents.“ (MILNE, Peter. *Tarski on Truth and Its Definition*. s. 1–2.)

²³ TARSKI, Alfred. *Sémantická koncepce pravdy a základy sémantiky*. s. 137.

²⁴ Viz KIRKHAM, Richard L. *Theories of Truth: A Critical Introduction*. s. 56–57.

²⁵ Tamtéž, s. 359.

²⁶ TARSKI, Alfred. *Sémantická koncepce pravdy a základy sémantiky*. s. 137.

1.3 Závislost pravdy na jazyku

Z Tarského rozhodnutí aplikovat pojem „pravdivý“ na věty vyvstávají jisté velmi závažné konsekvence, z nichž jednou z nejdůležitějších je nutnost relativizace tohoto pojmu vzhledem k jazyku. Nyní se pokusím dokázat, že není obtížné zdůvodnit požadavek na relativizaci pravdy vzhledem k jazyku. Uvažte následující příklad. Pokud je někde napsaná věta „horká čokoláda je tekutina“, pak tuto větu můžeme považovat za pravdivou větu v češtině, nepravdivou větu ve slovenštině a za nesmyslnou větu v prakticky jakémkoliv jiném jazyku. Tento jednoduchý příklad dokládá, že „jeden a týž výraz, jenž je pravdivou větou v jednom jazyce, může být nepravdivou nebo nesmyslnou větou v jazyce jiném“.²⁷ To je důvod, proč vždy, když je termín „pravdivý“ vztahován na věty, byl tento termín relativizován vzhledem k jazyku, ve kterém je formulována věta, které je připisována pravdivost nebo nepravdivost.

Dále je nutné si uvědomit, že důležitou roli při připisování pravdivosti libovolné větě hraje temporální faktor, respektive diachronie jazyka. To si dobře uvědomují např. lingvisté — konkrétně ti, kteří se zabývají tzv. diachronní lingvistikou. Ta se zabývá studiem jazyků s ohledem k jejich postupnému vývoji v průběhu času. Do jazyka se dostávají slova nová — tzv. computerové revoluci vděčíme za obohacení slovníku češtiny o slovo „harddisk“ —, význam již existujících slov se mění — příkladem budiž slovo „myš“ —, a jiná naopak z jazyka (alespoň z běžně používané slovní zásoby) mizí — např. disketa. Kromě toho je zapotřebí brát v úvahu i genezi a vývoj celých jazyků. Zatímco některé jazyky mizí a stávají se tak mrtvými jazyky — dokladem může být mykénština —, jiné jazyky zase vznikají — jako např. esperanto.

Logikové a filosofové si často vytvářejí vlastní „jazyky“ pouze za účelem dokázání vlastních tezí. Prakticky vždy jde buď o fragmenty jazyka, nebo o jazyky málo obsáhlé. Tyto jazyky jsou občas založeny na již existujících jazycích, ale občas jsou to jazyky diametrálně odlišné ode všech známých jazyků. Z tohoto si lze dovodit, že zohlednění diachronie jazyků není dostatečné. Je totiž myslitelné, že by mohly být i jiné jazyky, jejichž výrazy a jejich významy (stejně jako pravidla jejich syntaxe) by byly odlišné od všech diachronně uvažovaných jazyků. Je tak zapotřebí brát v úvahu nejenom reálně existující jazyky, ale zároveň s nimi všechny možné jazyky.

Z výše uvedeného je jasné, že pojem pravdy aplikovaný na věty je vždy nezbytně nutné relativizovat vzhledem k jazyku, ve kterém se vyskytuje (a který zkoumáme). Tímto

²⁷ Tamtéž.

však rozhodně není vyčerpáno všechno, co jsem chtěl o problematice relativizace k jazyku v této práci uvést. Nadále se tomuto tématu budu věnovat v souvislosti s T-schématem v kapitole 2.2.

1.4 Kritérium formální správnosti

Tarski klade na definici pravdy, o kterou mu jde, dva požadavky. Nyní se budu zabývat prvním z nich — požadavkem formální správnosti. Tarski sám se tomuto požadavku nikdy nevěnoval dostatečně podrobně. Toto tvrzení podporuje např. Milne:

Tarského druhou podmínkou kladenou na definice pravdy je, že musí být *formálně správné*, tuto podmínku občas vyjadřuje jako *metodologickou správnost*. Tato podmínka je poněkud méně jasně specifikována než podmínka materiální adekvátnosti, ale (kromě omezení toho, co může patřit do metajazyka) vyžaduje, aby definice odpovídaly standardním podmínkám kladeným na definice.²⁸

Bohužel ani v sekundární literatuře není situace o mnoho lepší. Jak Tarski, tak ostatní filosofové a logikové dávají ve svých pracích přednost spíše druhému Tarského požadavku — materiální adekvátnosti —, který proberu v následující kapitole. Tarski sám prohlašuje o formální správnosti následující:

Musíme tedy specifikovat slova či pojmy, jichž chceme k definování pravdy použít, a také musíme předložit formální pravidla, kterým by měla definice vyhovovat. Řečeno obecněji, musíme popsat formální strukturu jazyka, ve kterém bude definice podána.²⁹

Než však přejdu k dalším aspektům Tarského sémantické koncepce pravdy, hodlám se zmínit o dalších obecných hlediscích požadavku formální správnosti. Při konstrukci své definice nehodlá Tarski zavrhnout žádné běžně akceptované logické zákony a principy. Jedním z logických principů, který hraje velmi důležitou roli v Tarského koncepci je princip bivalence — podle něho má věta vždy jednu z pravdivostních hodnot „pravda“, nebo „nepravda“. Tarského přijetí principu bivalence bývá kritizováno, jak bude ukázáno v kapitole 10, věnované Kripkeho kritice Tarského pravdivostního predikátu.

²⁸ „Tarski’s second constraint on definitions of truth is that they must be *formally correct*, a condition he sometimes renders as *methodological correctness*. This is somewhat less clearly specified than the material adequacy condition but, in addition to limits on what can belong to the metalanguage, requires that the definitions conform to standard constraints imposed on definitions.“ (MILNE, Peter. Tarski on Truth and Its Definition. s. 4.)

²⁹ TARSKI, Alfred. Sémantická koncepce pravdy a základy sémantiky. s. 136–137.

Cílem kritéria formální správnosti je, aby z Tarského definice nebylo možné odvodit nepravdivá tvrzení (což úzce souvisí s požadavkem materiální adekvátnosti probíraným v následující kapitole), popř. nějaký paradox. Tarski se zabýval především možností vytvoření paradoxu lháře.³⁰ Paradox plodící větu lze formulovat např. následovně:

(LP) Věta (LP) je nepravdivá.³¹

Jak jsem uvedl dříve v této kapitole, pro větu (LP) musí platit normální logická pravidla. U jiných vět by to samozřejmě nepředstavovalo žádný problém, ale v případě věty (LP) dochází ke vzniku paradoxu, který lze schematicky znázornit takto:

- (1) (LP) je věta
- (2) Každá věta je buď pravdivá, nebo nepravdivá, ale nikdy nemůže být pravdivá a nepravdivá zároveň
- (3) (LP) je tedy buď pravdivá, nebo nepravdivá, ale nikoli obojí zároveň
- (4) Je-li (LP) pravdivá, pak je pravdivé to, co říká; (LP) je tedy nepravdivá
- (5) Je-li (LP) nepravdivá, pak není to, co říká, pravda; (LP) je tedy pravdivá
- (6) (LP) je tedy pravdivá i nepravdivá zároveň — podle (4) a (5)
- (7) (6) je ale zjevně v rozporu s (2) a (3) a tedy vzniká paradox³²

Zde je jasně vidět, že paradoxní situace spočívá v současné pravdivosti a nepravdivosti věty (LP). Pokud by se mělo zkoumání vzniku paradoxu lháře a možností jeho vytvoření omezit pouze na předpoklady obsažené v tomto schématu, pak by mohl leckdo nabýt dojmu, že vzniku paradoxu by bylo možné zabránit odmítnutím nebo pozměněním některého z nich. S tímto by však Tarski nesouhlasil, neboť v tomto schématu nejsou zohledněny všechny možné předpoklady, které mohou mít za následek možnost vzniku paradoxu lháře. Těmto předpokladům bude věnována kapitola 3.2 a kapitoly následující. Tam bude ukázán Tarského přístup k řešení paradoxu lháře. Toto řešení však bývá vystavováno kritice založené na upravené variantě paradoxu lháře — tzv. mstivém lháři. Stejně tak bývá Tarského práce napadána pomocí Jourdainova a Buridanova

³⁰ V částech této práce věnovaných variantám paradoxu lháře budu přebírat terminologii, kterou při zkoumání paradoxů používá dr. Raclavský. V souvislosti s paradoxem lháře půjde především o termíny *autoreferenční*, *heteroreferenční* a *(ne)kvantifikační*, kterých je využito pro klasifikaci jednotlivých druhů paradox plodících vět.

³¹ V tomto případě se jedná o autoreferenční nekvantifikační variantu paradoxu lháře. Autoreferenční varianty paradoxu lháře jsou založeny na schopnosti paradox plodící věty odkazovat k sobě samé, respektive připsovat pravdivostní hodnotu sobě samé. O nekvantifikačních variantách paradoxu lháře mluvíme tehdy, když se v paradox plodící větě nevyskytují žádné kvantifikátory.

Autoreferenční nekvantifikační variantu paradoxu lháře lze formulovat i poněkud neformálněji, bez použití jména dané věty. Paradox plodící věta v tomto případě zní: „Tato věta je nepravdivá.“

³² Srovnej KIRKHAM, Richard L. *Theories of Truth: A Critical Introduction*. s. 272–273.

paradoxu, což jsou taktéž speciální varianty paradoxu lháře. Obě tyto kritiky budou objasněny v kapitolách 6.1 a 6.2 v kritické části této práce.

1.5 Kritérium materiální adekvátnosti

Zde je třeba navázat na to, co bylo o Tarském a jeho přístupu k problému pravdivosti uvedeno již dříve. Především je třeba znovu uvést, že Tarskému nešlo o analýzu termínu hovorového jazyka „pravdivý“ používaného v každodenní konverzaci. Dále je zapotřebí vzít v úvahu, že Tarski chtěl, aby jeho definice pravdy „naplňovala intuice, jež se pojí s klasickým aristotelským pojetím pravdy“,³³ o kterém jsem již pojednal v první kapitole této práce. Přesnější vymezení Tarského cíle a metody lze najít u Soamese:

[P]odle Tarského a mnoha ostatních je náš běžný pojem pravdy defektní právě proto, že jeho neomezenost dává vzniknout paradoxům. Analýza pravdy, kterou se Tarski snaží poskytnout, je taková, která by eliminovala tento údajný defekt, zatímco by zachovala hlavní a užitečné prvky tohoto pojmu. Ve skutečnosti se Tarski nesnažil specifikovat, jak termín „pravdivý“ skutečně chápeme, ale jak by měl být pochopen, pokud má fungovat v našich logických, matematických a vědeckých teoriích způsobem normálně zamýšleným.

Analýzy tohoto druhu jsou často zvány „explikace“.³⁴

Tato Tarského snaha by šla vyjádřit za pomoci článku Rudolfa Carnapa O explikaci následovně: „Řekneme-li, že daný předvědecký pojem³⁵ je třeba přetransformovat do exaktního pojmu, znamená to ovšem, že je nutno zavést přesný pojem,³⁶ který odpovídá danému pojmu.“³⁷ Roli explikanda nemůže v Tarského koncepci hrát termín „pravdivý“ převzatý z hovorového jazyka. To by totiž bylo zcela nevyhovující zadání explikanda. Přílišná vágnost zadání explikanda by totiž klidně mohla mít za následek, že by se explikát minulo účinkem, jelikož by bylo explikováno něco úplně jiného, než o co Tarskému ve skutečnosti šlo.

³³ TARSKI, Alfred. *Sémantická koncepce pravdy a základy sémantiky*. s. 138.

³⁴ „[A]ccording to Tarski and many others, our ordinary notion of truth is defective precisely because its unrestrictedness gives rise to paradox. The sort of analysis of truth that Tarski attempted to provide is one that would eliminate this alleged defect while preserving the central and useful features of the notion. In effect, Tarski tried to specify not how ‘true’ is actually understood but how it ought to be understood if it is to function in our logical, mathematical, and scientific theories in the ways normally intended. Analyses of this kind are often called “explications”.“ (SOAMES, Scott. *Understanding Truth*. Oxford; New York: Oxford University Press, 1999. s. 99.)

³⁵ Explikandum.

³⁶ Explikát.

³⁷ CARNAP, Rudolf. *Problémy jazyka vědy: (výbor prací Rudolfa Carnapa)*. Praha: Svoboda, 1968. s. 170.

Na počátku obou prací se Tarski přihlásil ke korespondenční teorii pravdy³⁸ a intuicím s ní spojovaným. Problém pravdivosti tak dostává docela zřetelné obrysy a je jasné, co má být náplní zbytku obou Tarského zmiňovaných prací. Je ale nezbytné zmínit, že i v tomto ohledu panují neshody ohledně Tarského koncepce. I přes jeho explicitní přihlášení se ke korespondenční teorii pravdy si někteří filosofové kladou otázku, jestli je Tarského koncepci pravdy možné považovat za korespondenční.³⁹

Cílem Tarského snažení nebylo vytvoření nové teorie pravdy, ale podání adekvátní explikace korespondenční teorie pravdy, přičemž jejím nejjasnějším příkladem byla klasická definice. Tato Aristotelova formulace jasně určuje dva případy, kdy je to, co říkáme, pravdivé — totiž pokud o tom, co je, prohlásíme, že to je; a pokud o tom, co není, řekneme, že to není. Zdá se, že Aristotelés by (s ohledem na svoji formulaci) souhlasil s tvrzením, že věta „sníh je bílý“ může být pravdivá pouze tehdy, když sníh je bílý — tedy pouze tehdy, když platí následující ekvivalence:

(5) Věta „sníh je bílý“ je pravdivá tehdy a jen tehdy, když sníh je bílý.⁴⁰

Tato ekvivalence musí podle Tarského být důsledkem definice pravdy. Je snadné pochopit proč — kdyby z Tarského definice pravdy neplynula tato věta, která je intuitivně pravdivá, pak by to nemohla být skutečně obecná definice pravdy, jelikož by existovaly pravdivé věty, které by z této definice nebyly odvoditelné.

³⁸ Tarski se explicitně přihlásil ke klasické variantě korespondenční teorie pravdy, jak bylo ukázáno již v první kapitole této práce. Mezi interpretátory Tarského jsou však vedeny diskuse, jestli Tarski sám chtěl, aby byla jeho koncepce korespondenční. A pokud ano, jestli se mu tento úkol podařilo splnit. Srovnej KIRKHAM, Richard L. *Theories of Truth: A Critical Introduction*. s. 170–173.

³⁹ Srovnej LYNCH, Michael B. *The Nature of Truth: Classic and Contemporary Perspectives*. s. 327.

⁴⁰ TARSKI, Alfred. *Sémantická koncepce pravdy a základy sémantiky*. s. 139.

2 Konvence T

2.1 Konvence T a T-schéma

K tomu, abychom věděli, jaké všechny věty by měly plynout z definice pravdy je zapotřebí zobecnit větu (5) z předchozí kapitoly, což lze provést poměrně jednoduše. Stačí si povšimnout jednoho velmi důležitého detailu. V ekvivalenci (5) se vyskytuje výraz „sníh je bílý“ dvakrát — v prvním výskytu je tento výraz uzavřený do uvozovek, ve druhém však nikoliv. Zatímco na pravé straně se jedná o výraz samotný, na levé straně věty se vyskytuje jeho jméno.

Výskyt jména výrazu ve výše uvedené ekvivalenci (5) má zcela klíčový význam. Tarski o výskytu jmen v ekvivalencích jako je (5) prohlásil: „chceme-li o nějaké větě např. říci, že je pravdivá, musíme použít jméno této věty, a nikoli tuto větu samotnou.“⁴¹ Je otázkou, nakolik výstižná je tato formulace, jelikož ať už se mluví o nějaké větě nebo o nějakém objektu, vždy je automaticky používáno jméno této věty nebo objektu. Např. ve výpovědi o větě, kterou kdosi vyslovil, se nemůže vyskytovat ona věta sama, ale pouze její jméno. Toto pravidlo lze vztáhnout prakticky na všechny předměty — jestliže např. někdo vysloví libovolnou větu o Slunci, vždy se v této větě objeví pouze jméno Slunce, ale Slunce samo, tedy ona olbřímí koule žhavých plynů, se v ní samozřejmě objevit nemůže. Samotnému tvoření jmen vět a jejich distinkci je věnována následující kapitola.

Nyní je zapotřebí vytvořit obecné schéma, ze kterého by byla výše uvedená ekvivalence odvoditelná. Toto schéma vznikne, když za větu, o které chceme prohlásit, že je pravdivá, dosadíme písmeno „*p*“ a za jméno této věty dosadíme písmeno „*X*“. Výsledkem zobecnění je toto schéma:

(T) *X* je pravdivá tehdy a jen tehdy, když *p*.⁴²

Pro toto schéma⁴³ se v literatuře zažilo označení T-schéma. Každou ekvivalenci vzniklou dosazením do T-schématu je zvykem označovat jako ekvivalenci tvaru (T), nebo jako T-větu.

⁴¹ Tamtéž.

⁴² Tamtéž, s. 140.

⁴³ Výraz (T) není větou, ale větným schématem, věta se z něj stává až dosazením (metajazykového překladu) libovolné věty za *p* a jejího jména za *X*.

Roli T-schématu v rámci Tarského koncepce lze lépe pochopit po seznámení se s obsahem tzv. konvence T:

KONVENCE T. *Formálně správná definice symbolu ,Tr' formulovaná v metajazyku bude nazývána adekvátní definicí pravdy, pokud má následující důsledky:*

(α) *všechny věty získané z výrazu , $x \in Tr$ tehdy a jen tehdy, když p' nahrazením symbolu , x' strukturálně-deskriptivním jménem jakékoliv věty zkoumaného jazyka a nahrazením symbolu , p' výrazem tvořícím překlad této věty do metajazyka;*

(β) *věty ,pro jakékoli x , jestliže $x \in Tr$, pak $x \in S'$ (jinými slovy , $Tr \subseteq S'$).'⁴⁴*

Klíčovou roli hraje především důsledek (α), podle kterého musí mít každá adekvátní definice pravdy za následek všechny ekvivalence tvaru (T) získané dosazením vět zkoumaného jazyka do T-schématu. Zde je zapotřebí si uvědomit, že toto je požadavek, který musí splnit každá teorie pravdy, ale rozhodně se zde nejedná o teorii pravdy samu. T-schéma nemůže být teorií pravdy právě proto, že jde jen o větné schéma.

Ačkoliv T-schéma není definicí pravdy, mohl by padnout návrh, že definicí pravdy bude konjunkce všech ekvivalencí tvaru (T), které je možné vytvořit pro daný zkoumaný jazyk. Jak dodává Tarski, každou jednotlivou ekvivalenci tvaru (T) „lze považovat za dílčí definici pravdy, jenž vysvětluje, v čem spočívá pravdivost této jedné jednotlivé věty“.⁴⁵ Konjunkce všech ekvivalencí tvaru (T) by opravdu byla definicí pravdy, nicméně toto řešení naráží na jisté metodologické problémy. Pokud by byl zkoumaný jazyk silně omezený v tom smyslu, že by připouštěl konstrukci pouze omezeného množství vět, nepředstavovala by tvorba definice pro tento jazyk výraznější problémy. Ty by však nastaly v případě, pokud by bylo možné ve zkoumaném jazyku zkonstruovat nekonečné množství vět. Definice pravdy pro takový jazyk by pak musela být konjunkcí nekonečného množství vět, což není optimální řešení. Tarského přístup k tvorbě definice pravdy je proto odlišný a bude představen v kapitole 4.

⁴⁴ „CONVENTION T. A formally correct definition of the symbol 'Tr', formulated in the metalanguage, will be called an adequate definition of truth if it has the following consequences:

(α) all sentences which are obtained from the expression ' $x \in Tr$ if and only if p' ' by substituting for the symbol ' x ' a structural-descriptive name of any sentence of the language in question and for the symbol ' p' ' the expression which forms the translation of this sentence into the metalanguage;

(β) the sentence 'for any x , if $s \in Tr$ then $x \in S'$ ' (in other words ' $Tr \subseteq S'$ ').“ (TARSKI, Alfred. The Concept of Truth in Formalized Languages. s. 187–188.)

⁴⁵ TARSKI, Alfred. Sémantická koncepce pravdy a základy sémantiky. s. 140.

2.2 Rozšíření T-schématu

V předchozí kapitole jsem představil T-schéma tak, jak ho ve svých pracích používá Tarski. Jeho verze je však poněkud „chudá“, jelikož obsahuje jako pravdivostní predikát pouze „pravdivý“, popř. „Tr“ ve formálněji laděném článku [CTFL]. To je poněkud zvláštní, jelikož Tarski explicitně uvádí, že termín „pravdivý“ musí být vždy relativizován vzhledem k jazyku. A nejideálnějším způsobem, jak dát tuto relativizaci explicitně najevo, se jeví její zavedení do T-schématu, což však Tarski sám neprovedl. Šťastným se v této souvislosti jeví, že na relativizaci pravdivostního predikátu v T-schématu nezapomněli Tarského komentátoři — jak uvádí Boisvert: „[J]estliže je pravdivostní predikát aplikován na věty, pak je nezbytně relativizován vzhledem k jazyku, neboť stejný řetězec symbolů může mít různé významy v různých jazycích.“⁴⁶

Používat v T-schématu pravdivostní predikát s explicitním zohledněním jazyka, na který je tento predikát aplikovatelný, je v literatuře běžné. Příkladem T-schématu, ve kterém je závislost na jazyku zohledněna můžeme najít např. u Scotta Soamese:

Schéma T. X je T v L tehdy a jen tehdy, když P.⁴⁷

Prvky tohoto schématu jsou následující: X je jméno věty, zatímco P je jejím překladem do metajazyka, T je pravdivostní predikát a L je jazyk, na jehož věty je pravdivostní predikát aplikovatelný.⁴⁸ Zde bych si dovolil podotknout, že T a L se většinou vyskytují spojené do jednoho celku. V literatuře lze najít i mnohé jiné varianty notace relativizovaného pravdivostního predikátu, jako např.: „pravdivý-v-L“, „pravdivý_L“, „Tr_L“ a podobně.

Jelikož Tarského zajímala pouze možnost konstrukce definice pravdy pro formalizované jazyky, mohla by se jevit relativizace pravdivostního predikátu vzhledem k jazyku dostatečnou. Mnozí autoři však počítají i s možností aplikace Tarského koncepce na hovorové jazyky, přičemž relativizace pravdivostního predikátu pouze k jazyku se nezdá být pro jejich účely dostatečná. Proto bývá pravdivostní predikát vztahován i k jiným hlediskům.

⁴⁶ „[T]he truth predicate is necessarily relativized to a language if it is applied to sentences, since the same string of symbols can have different meanings in different languages.“ (BOISVERT, Daniel R. *The Trouble with Harrison's 'The Trouble with Tarski'*. *The Philosophical Quarterly*. July 1999, vol. 49, no. 196, s. 377.)

⁴⁷ „Schema T. X is T in L iff s.“ (SOAMES, Scott. *Understanding Truth*. s. 68.)

⁴⁸ L je tedy jazyk, pro který je definice pravdy konstruována.

Boisvert ve své kritice Harrisona používá rozšířenou verzi T-schématu, aby mohl Tarského koncepci pravdy aplikovat na příklady obsahující věty hovorového jazyka. Jeho varianta T-schématu vypadá následovně:

TN. S je pravdivá $_{\{u,t\}}$ v L tehdy a jen tehdy, když p .⁴⁹

V tomto schématu představuje S jméno věty, p její metajazykový překlad a L objektový jazyk, pro který je definována teorie pravdy. Výraznější změnou oproti předchozí rozšířené variantě T-schématu představuje přidání parametrů u a t k pravdivostnímu predikátu pravdivá $_{\{u,t\}}$. Oborem u jsou mluvčí a oborem t jsou momenty v čase. Boisvert tak kromě k jazyku relativizuje pravdivostní predikát pro přirozený jazyk ještě k mluvčímu, který větu pronesl, a k času, kdy byla věta pronesena. Soames se zmiňuje o koncepci, v níž je pravda relativizována vzhledem k okolnostem, které mohou být chápány jako částečné stavy světa, nebo jako možné světy. Pravdivostní predikát v ní má podobu „pravdivý v C “.⁵⁰ Ale tímto se již příliš vzdalují od samotné Tarského koncepcie, pročež radši přejdu k dalšímu tématu.

2.3 Dva druhy jmen vět

V předchozí kapitole jsem uvedl, že v T-schématu se vyskytuje jednak věta a pak i její jméno. Tuto kapitolu tedy věnuji problematice vytvoření jména věty. V [CTFL]⁵¹ Tarski uvádí a definuje dva různé způsoby, jak vytvořit jméno věty,⁵² které lze posléze dosadit do T-schématu.

Prvním druhem jsou citační jména.⁵³ Vytvoření jména věty proběhne jednoduše uzavřením dané věty do uvozovek — jméno se tedy bude skládat pouze z věty samotné a uvozovek na obou jejích stranách. Tarski o citačních jménech dále uvádí:

S citačními jmény může být zacházeno jako s jednotlivými slovy jazyka, tudíž jako se syntakticky jednoduchými výrazy. Jednotlivé konstituenty těchto jmen — uvozovky a výrazy stojící mezi nimi — plní stejnou funkci jako písmena a komplexy následujících písmen v jediném slově. [...] Každé citační jméno je tedy konstantním individuálním jménem určitého výrazu

⁴⁹ „TN. S is true $_{\{u,t\}}$ in L if and only if s .“ (BOISVERT, Daniel R. The Trouble with Harrison's 'The Trouble with Tarski'. s. 378.)

⁵⁰ Srovnej SOAMES, Scott. *Understanding Truth*. s. 137–138.

⁵¹ Viz TARSKI, Alfred. The Concept of Truth in Formalized Languages. s. 156–158.

⁵² Jen pro úplnost uvádím, že těmito způsoby se dají vytvořit jména libovolných výrazů, nejenom vět.

⁵³ Tarského termín „quotation-mark names“ pro účely této práce překládám jako „citační jména“.

(výrazu uzavřeného do uvozovek) a ve skutečnosti jménem stejné povahy, jako je vlastní jméno člověka.⁵⁴

Citační jména jsou jednoduchá právě v tom smyslu, že každé slouží jako individuální jméno daného výrazu. Vnitřní struktura citačního jména je sice složená, zvláště jedná-li se o jméno věty, ale vždy je s ním zacházeno jako s jediným slovem/jménem. Tarského přirovnání k vlastním jménům lidí je velmi trefné — ostatně jméno „Bertrand Arthur William, 3. hrabě Russell“ má rozhodně složenou vnitřní strukturu, ale musíme s ním zacházet jako s jediným slovem/jménem určité osoby.⁵⁵

V praxi by použití citačních jmen při dosazení do T-schématu vypadalo následovně:

(5) Věta „sníh je bílý“ je pravdivá tehdy a jen tehdy, když sníh je bílý.⁵⁶

Dalším druhem jmen vět jsou strukturálně-deskriptivní jména. Postup jejich vytvoření je poněkud složitější než v předchozím případě. Tato jména necitují výraz jimi denotovaný, ale jistým způsobem ho popisují. K vytvoření strukturálně-deskriptivního jména věty je zapotřebí mít jména pro všechna písmena jazyka, do kterého daná věta náleží, a také jména pro všechny ostatní v něm použité symboly (interpunkční znaménka, matematické a logické symboly a podobně) — tato jména ale sama nemohou být citačními jmény. Samotné vytvoření je možné připodobnit ke skládání puzzle — vezmeme jména pro všechny symboly ve větě se vyskytující a poskládáme je tak, aby uspořádání jmen symbolů odpovídalo uspořádání symbolů jimi denotovaných. Při dosazení strukturálně-deskriptivních jmen do T-schématu dostaneme následující ekvivalenci tvaru (T):

(6) Věta skládající se ze tří slov, z nichž první má čtyři písmena — Es, En, Dlouhé-I a Há —, druhé má dvě písmena — Jé a É — a třetí ze čtyř písmen — Bé, Dlouhé-I, El a Dlouhé-Y, je pravdivá tehdy a jen tehdy, když sníh je bílý.⁵⁷

⁵⁴ „Quotation-mark names may be treated like single words of a language, and thus like syntactically simple expressions. The single constituents of these names — the quotation marks and the expressions standing between them — fulfil the same function as the letters and complexes of successive letters in single words. [...] Every quotation-mark name is then a constant individual name of a definite expression (the expression enclosed by the quotation marks) and in fact a name of the same nature as the proper name of a man.“ (TARSKI, Alfred. *The Concept of Truth in Formalized Languages*. s. 159.)

⁵⁵ S tímto pojetím citačních jmen se pojí jisté potíže, na které poukazuje Soames v SOAMES, Scott. *Understanding Truth*. s. 86–87.

⁵⁶ TARSKI, Alfred. *Sémantická koncepce pravdy a základy sémantiky*. s. 139.

⁵⁷ Toto ale není jediný způsob, jak lze strukturálně-deskriptivní jméno vytvořit. V [SCT] Tarski používá jako jména jednotlivých písmen jejich pořadí v abecedě, srovnej tamtéž, s. 140.

Oba druhy pojmenování v této kapitole probírané je samozřejmě možné mezi sebou zaměňovat — jména uzavřená do uvozovek je možné nahradit jmény strukturálně-deskriptivními a naopak.

3 Tarského dělení jazyků

3.1 Jazyky se specifikovanou strukturou, formalizované jazyky

V této kapitole se budu zabývat jazyky, pro které chtěl Tarski podat definici pravdivosti. Tarski explicitně uvádí: „Problém definice pravdy nabývá přesného významu a může být rigorózně vyřešen jen pro jazyky, jejichž struktura byla exaktně specifikována.“⁵⁸ Bude tedy záhodno definovat jazyky se specifikovanou strukturou, což bude nejjednodušeji proveditelné tak, že uvedu kritéria, kterým musí tyto jazyky vyhovovat:

- i. Třída všech slov a všech výrazů považovaných v daném jazyku za smysluplné musí být jednoznačně popsána
- ii. Musí být uvedeny všechny primitivní (nedefinované) termíny a zároveň musí být dána pravidla definice umožňující zavádění definovaných termínů do jazyka
- iii. Musí být stanovena kritéria, podle kterých lze z třídy výrazů daného jazyka vyčlenit věty a následně musí být stanovena pravidla určující podmínky, za nichž je možné větu tvrdit
- iv. Aby bylo možné odvozovat nové věty, je zapotřebí určit axiomy (primitivní věty) daného jazyka a pravidla vyplývání (pravidla inference), pomocí kterých z axiomů odvodíme nové věty, které budou nazývány teorémy (tedy věty dokazatelné v daném jazyku)⁵⁹

Jedinými jazyky se specifikovanou strukturou byly v době, kdy psal Tarski svoje články, formalizované jazyky různých systémů deduktivní logiky, jejichž účelem bylo zkoumání deduktivních věd. Tarského cílem bylo podat definici pojmu „pravdivý“ právě pro formalizované jazyky. Proč tomu tak bylo, odhalí následující úryvek:

[T]arski podporoval fyzikalismus, neboli názor že jakýkoliv pojem stojící za zmínku může být definován kompletně v termínech odkazujících pouze k fyzikálním a matematickým entitám. Důsledkem tohoto si Tarski stanovil úkol definovat pravdu tímto způsobem pro jisté formální jazyky, nebo jazyky, které měly být adekvátní pro fyzikální vědy a matematiku.⁶⁰

Formalizované jazyky jsou všechny takové jazyky, v nichž mohou být tvrzeny pouze teorémy. Ve formalizovaném jazyku tak můžeme tvrdit pouze ty věty, které jsme schopni odvodit za pomoci usuzovacích pravidel z axiomů daného jazyka — tedy pouze

⁵⁸ Tamtéž s. 143.

⁵⁹ Volně podle TARSKI, Alfred. *Sémantická koncepce pravdy a základy sémantiky*. s. 142–143.

⁶⁰ „[T]arski subscribed to physicalism, or the view that any concept worth having can be defined completely in terms that refer only to physical and mathematical entities. As a result, Tarski set himself the task to define truth in this way for certain formal languages, or languages thought to be adequate for the physical sciences and mathematics.“ (LYNCH, Michael B. *The Nature of Truth: Classic and Contemporary Perspectives*. s. 324.)

dokazatelné věty. V takovém jazyku hraje hlavní roli forma jeho výrazů — význam každého výrazu tohoto jazyka je jednoznačně dán jeho formou. Podobně, jako kritéria pro jazyky se specifikovanou strukturou uvádí Tarski na počátku §2 [CTFL]⁶¹ kritéria pro formalizované jazyky, avšak vzhledem k jejich veliké vzájemné podobnosti je zde nebude uvádět.

Je vhodné dodat, že formalizované jazyky a jazyky se specifikovanou strukturou nejsou vzájemně zaměnitelné termíny. Jak totiž Tarski uvádí, lze si „představit jazyky, jež by měly exaktně specifikovanou struktur, aniž by byly formalizované“.⁶² V těchto jazycích není tvrditelnost vět závislá výlučně na jejich formální stránce, ale může být závislá i na mimojazykových faktorech. Z toho a z definic uvedených výše vyplývá, že takový jazyk by nemohl být formalizovaný. Roli těchto jazyků Tarski spatřoval v tom, že pokud by byly vytvořeny, mohly by teoreticky převzít roli hovorových jazyků při vědeckém diskursu.

Z výše uvedeného je jasné, že Tarskému šlo o podání definice termínu „pravdivý“ pro formalizované jazyky, respektive pro jazyky se specifikovanou strukturou. Na tomto místě se vnucuje otázka, jestli existuje možnost vytvořit takovou definici i pro hovorové jazyky a jestli Tarski o této možnosti uvažoval. Tarski se samozřejmě vyjádřil i k této otázce, já se jí budu zabývat v následujících kapitolách.

3.2 Sémanticky uzavřené jazyky

Než se pokusím pojednat o Tarského postoji k hovorovým jazykům, připravím si teoretické základy pojednáním o jazycích, které Tarski ve svých zkoumáních zavrhuje jakožto nekonzistentní. Ostatně i Tarski se k problematice hovorových jazyků a možnosti vytvoření definice termínu „pravdivý“ pro ně dostává jaksí oklikou přes takzvané sémanticky uzavřené jazyky. Budu tedy následovat jeho vzor a rovněž nejdříve pojednám o sémanticky uzavřených jazycích.

Sémanticky uzavřený jazyk je dle Tarského takový jazyk, který vyhovuje prvním dvěma z následujících předpokladů:

- a) Tento jazyk obsahuje kromě vlastních výrazů také jména těchto výrazů (nebo jiný způsob umožňující mluvit o větách tohoto jazyka) a sémantické termíny odkazující k větám tohoto jazyka⁶³

⁶¹ Viz TARSKI, Alfred. *The Concept of Truth in Formalized Languages*. s. 166.

⁶² TARSKI, Alfred. *Sémantická koncepce pravdy a základy sémantiky*. s. 143.

⁶³ Srovnej SOAMES, Scott. *Understanding Truth*. s. 67.

- b) V tomto jazyku je možné tvrdit všechny věty, vymezuující adekvátní použití těchto sémantických termínů⁶⁴
- c) V tomto jazyku platí obvyklé logické zákony

Co se stane, pokud daný jazyk splňuje první dva předpoklady a zároveň vyhovuje předpokladu (c), bylo ukázáno na konci kapitoly věnované formální správnosti. V takovém jazyku je totiž možné formulovat paradox lháře. Nyní se znovu zaměřím na paradox plodící větu (LP) a následně poukážu, jak výše zmíněné předpoklady vedou ke vzniku paradoxu:

(LP) Věta (LP) je nepravdivá.

Podle předpokladu (a) mají věty ve zkoumaném jazyku možnost odkazovat samy k sobě a navíc mohou obsahovat na sebe aplikovatelné sémantické predikáty. Věta (LP) této podmínce nepochybně vyhovuje. Obsahuje totiž svoje vlastní jméno, díky kterému odkazuje sama k sobě. Ovšem jméno není to jediné, za pomoci čeho by k sobě mohla odkazovat. Ostatně paradox by vzniknul i v případě, že by paradox plodící věta zněla: „Tato věta je nepravdivá.“ I ta je totiž schopná vypovídat sama o sobě. Dále věta (LP) obsahuje sémantický termín „nepravdivý“, který aplikuje sama na sebe, čímž naplňuje i druhou polovinu předpokladu (a). Podle předpokladu (b) musí být věta (LP) příkladem adekvátního použití termínu „nepravdivý“. Tvrzení, že nějaká věta je nepravdivá, jsou v jazyku běžná a (LP) tak vyhovuje i tomuto předpokladu. A konečně vznik paradoxu je těsně vázán na bod (c), jelikož kdyby neplatily obvyklé logické zákony, jako je zákon vyloučeného třetího, ke vzniku paradoxu by nemuselo dojít.

Ještě než uvedu, který (nebo které) z výše zmíněných důvodů Tarski odmítá jako hlavní příčinu paradoxů, se krátce zmíním o jisté specifičnosti paradoxu lháře. Výše zmíněná varianta je totiž závislá na použití termínu „nepravdivý“, který symbolicky můžeme označit jako „ $\neg Tr$ “. Kdyby se totiž v paradox plodící větě objevil sémantický termín „pravdivý“ („ Tr “), vůbec by ke vzniku paradoxu nedošlo.

Podle Tarského není akceptovatelné, aby byl kvůli odstranění možnosti tvorby paradoxu odmítnut předpoklad (c). To by znamenalo vzdát se obvyklých logických zákonů, což je nepřijatelné. Předpoklady, které je zapotřebí odmítnout, aby nebylo v daném jazyku možné zkonstruovat antinomii lháře, jsou předpoklady (a) a (b). Aby mohl být zkoumaný jazyk konzistentní, pak podle Tarského nesmí obsahovat žádné prostředky, kterými by bylo možné odkazovat k jeho vlastním větám. Tarski tímto požadavkem

⁶⁴ Srovnej TARSKI, Alfred. Sémantická koncepce pravdy a základy sémantiky. s. 145.

z daného jazyka odstraňuje všechny autoreferenční věty, což mu bude také vytýkáno. Ve druhé části této práce představím kritiku tohoto postupu pocházející od Roberta L. Martina, podle které je odmítnutí všech autoreferenčních vět krokem špatným směrem a která nabízí jiný přístup k řešení sémantických paradoxů, než jaký nabízí Tarski.

Dále je nepřípustné, aby zkoumaný jazyk obsahoval sémantické termíny aplikovatelné na výrazy tohoto jazyka, což je důsledkem Tarského teorému, podle něhož je nemožné, aby byla definice pravdy vytvořena pouze s použitím pojmů objektového jazyka. Tomuto požadavku budu věnovat pozornost v kapitole 3.4 dedikované distinkci mezi objektovým jazykem a metajazykem.

3.3 Status hovorových jazyků

Nyní už mám dostatečné teoretické podklady k pojednání o Tarského postoji k hovorovým jazykům. V předminulé kapitole jsem uvedl, že Tarskému jde o podání definice pravdy pro jazyky se specifikovanou strukturou a následně jsem uvedl pravidla, kterým musí každý takový jazyk vyhovovat. V minulé kapitole jsem se zabýval sémanticky uzavřenými jazyky, tedy jazyky, pro které je nemožné zkonstruovat definici pravdy. Nyní se budu zabývat tím, co Tarski napsal o hovorových jazycích.

V předchozí kapitole jsem uvedl, že každý sémanticky uzavřený jazyk musí vyhovovat třem podmínkám. Každý jazyk, který těmto podmínkám opravdu vyhovuje, pak Tarski považuje za nekonzistentní. V souladu s podmínkami (a) a (b) je zapotřebí zmínit jeden důležitý rys hovorového jazyka — jeho univerzálnost.

můžeme tvrdit, že ‚pokud o něčem můžeme smysluplně mluvit, pak o tom můžeme mluvit v hovorovém jazyce‘. Pokud má být zachována univerzálnost každodenního jazyka, [...] musíme v něm, abychom byli konzistentní, připustit kromě jeho vět a dalších výrazů také jména těchto vět a výrazů, věty obsahující tato jména, stejně tak i sémantické výrazy jako ‚pravdivá věta‘, ‚jmenovat se‘, ‚denotovat‘ atd.⁶⁵

Úkolem hovorového jazyka je tak podle Tarského mluvit o všem, o čem se mluvit dá. Jelikož hovorový jazyk je primárním prostředkem komunikace mezi lidmi, musí být dostatečně expresivní, aby mohl svoji roli dobře plnit. To však znamená, že takový jazyk

⁶⁵ „it could be claimed that ‘if we can speak meaningfully about anything at all, we can also speak about it in colloquial language’. If we are to maintain this universality of everyday language [...] we must, to be consistent, admit into the language, in addition to its sentences and other expressions, also the names of these sentences and expressions, and sentences containing these names, as well as such semantic expressions as ‘true sentence’, ‘name’, ‘denote’, etc.“ (TARSKI, Alfred. *The Concept of Truth in Formalized Languages*. s. 164.)

bude muset (aby si zachoval svoji univerzálnost) mít i ty vlastnosti, které má sémanticky uzavřený jazyk. Z výše uvedeného citátu si lze odvodit, že hovorový jazyk vyhovuje podmínkám (a) a (b) charakterizujícím sémanticky uzavřený jazyk. Lze tedy očekávat, že ho Tarski označí jako nekonzistentní — on však poněkud překvapivě konstatuje:

Náš každodenní jazyk jistě není jazykem s exaktně specifikovanou strukturou. [...] Z tohoto důvodu nemá problém konzistence ve vztahu k tomuto jazyku přesný význam. V nejlepším případě se můžeme odvážit tvrdit, že jazyk, jehož struktura by byla exaktně specifikována a jenž by se našemu každodennímu jazyku podobal do takové míry, do jaké by to jen bylo možné, by byl nekonzistentní.⁶⁶

I přes toto tvrzení je poměrně běžně uznávaným faktem, že Tarski považoval hovorové jazyky za nekonzistentní. Jak je ale vidět na předchozích citátech, toto tvrzení není až tak jednoduché dokázat. Jedním z charakteristických rysů Tarského práce totiž je, že „byl výjimečně opatrný ohledně podávání přímých filosofických interpretací svého díla“.⁶⁷ To je jeden z důvodů, proč panují debaty ohledně některých bodů jeho díla. Problematika nekonzistence hovorových jazyků není výjimkou. Existuje skupina filosofů zastávajících názor, že Tarski opravdu považoval hovorové jazyky za nekonzistentní a tento názor napadají. Příkladem může být Hans Herzberger, jehož kritikou Tarského se budu zabývat ve druhé části této práce. Existují ovšem i filosofové, podle nichž je neoprávněné připisovat Tarskému tvrzení, že hovorové jazyky jsou nekonzistentní. Příkladem tohoto budiž diskuse mezi A. B. Levisonem⁶⁸ a Vitem Sinisim.⁶⁹

3.4 Objektový jazyk a hierarchie metajazyků

Aby bylo možné zkonstruovat uspokojivou definici pravdy pro daný jazyk, je zapotřebí nejdříve zavést jednu nezbytně nutnou distinkci — distinkci mezi objektovým jazykem a metajazykem. Objektový jazyk je jazyk, který zkoumáme a pro který se budeme snažit zkonstruovat definici pravdy — je to tedy jazyk, o kterém budeme mluvit. Pokud už tedy máme daný jazyk, o kterém budeme mluvit, je nasnadě, že budeme potřebovat jazyk, ve kterém budeme mluvit. Tím bude metajazyk — tedy jazyk, ve kterém budeme pronášet

⁶⁶ TARSKI, Alfred. Sémantická koncepce pravdy a základy sémantiky. s. 146.

⁶⁷ „he was extraordinarily cautious and careful in giving any direct philosophical interpretation of his work.“ (SUPPES, Patrick. Philosophical Implications of Tarski's Work. s. 80.)

⁶⁸ LEVISON, A. B. Logic, Language, and Consistency in Tarski's Theory of Truth. *Philosophy and Phenomenological Research*. March 1965, vol. 25, no. 3, s. 384–392.

⁶⁹ SINISI, Vito F. Tarski on the Inconsistency of Colloquial Language. *Philosophy and Phenomenological Research*. June 1967, vol. 27, no. 4, s. 537–541.

výpovědi o objektovém jazyku a ve kterém bude zkonstruována definice pravdy pro objektový jazyk.

Ačkoli hraje distinkce mezi objektovým jazykem a metajazykem velmi důležitou roli v Tarského koncepci, on sám není jejím tvůrcem. Jak uvádí Peter Milne, Tarski přebral tuto distinkci z tradice Lvovsko-varšavské školy. Za jejího autora je považován Stanisław Leśniewski, který s její pomocí argumentoval proti nedostatkům v *Principia Mathematica*.⁷⁰

Tarski uvádí o metajazyku následující: „[M]etajazyk obsahuje jak individuální jména, tak i překlady všech výrazů (konkrétně všech vět) studovaného jazyka.“⁷¹ Jelikož účelem metajazyka je hovořit o objektovém jazyku, musí být metajazyk dostatečně bohatý, aby v něm bylo možné odkazovat k větám zkoumaného objektového jazyka. Toho lze docílit tak, že metajazyk bude obsahovat prvky potřebné ke zkonstruování jmen všech výrazů objektového jazyka. Toto však samo o sobě nestačí. Dále je totiž nezbytně nutné, aby metajazyk obsahoval překlady všech výrazů objektového jazyka.⁷² Tyto překlady hrají důležitou roli při tvorbě ekvivalencí tvaru (T). Každá taková ekvivalence musí mít na levé straně jméno zkoumané věty v metajazyku, a na pravé straně adekvátní překlad věty označované tímto jménem do metajazyka.⁷³ Kdyby se v objektovém jazyku vyskytovala věta, jejíž jméno a překlad by se nenacházely v metajazyku, pak by nebylo možné o této větě v metajazyku nic vypovědět a zkoumání tohoto objektového jazyka by tedy nemohlo být nikdy úplné.

Výčet prvků metajazyka se ovšem nevyčerpává překlady všech výrazů objektového jazyka spolu s jejich jmény. Co všechno ještě metajazyk může a naopak nesmí obsahovat, je možné ilustrovat si na příkladu T-schématu.

(T) X je pravdivá tehdy a jen tehdy, když p .

⁷⁰ Srovnej MILNE, Peter. Tarski on Truth and Its Definition. s. 2.

⁷¹ „[T]he metalanguage contains both an individual name and a translation of every expression (and in particular of every sentence) of the language studied.“ TARSKI, Alfred. The Concept of Truth in Formalized Languages. s. 173.)

⁷² Srovnej tamtéž, s. 188.

⁷³ Toto je nutné brát v úvahu především při snahách o aplikaci Tarského koncepce na přirozené jazyky. Pokud si např. zvolíme jako objektový jazyk angličtinu, jako zkoumanou větu „snow is white“ a jako metajazyk češtinu, pak ekvivalenci tvaru (T) zkonstruujeme následovně:

Věta „snow is white“ je pravdivá v angličtině tehdy a jen tehdy, když sníh je bílý.

Některé body kritiky Jonathana Harrisona projednávané v kapitole 9 jsou založeny na špatném překladu do metajazyka. Závěrem musím zdůraznit, že Tarskému nešlo o definování pravdy pro hovorové jazyky a tato poznámka je tak formulována s ohledem na moderní snahy o aplikaci Tarského koncepce na hovorové jazyky.

Již bylo uvedeno, že metajazyk obsahuje překlady výrazů (věty) objektového jazyka — těm odpovídá v T-schématu „ p “ — a dále jejich jména — tedy X v T-schématu. Kromě toho je zapotřebí, aby byly v metajazyku obsaženy „termíny obecně logické povahy“, ⁷⁴ tedy termíny především z výrokové a predikátové logiky, jako je např. „tehdy a jen tehdy, když“.

Dalším prvkem metajazyka jsou sémantické termíny vztahující se k výrazům objektového jazyka, k jehož zkoumání a popisování daný metajazyk používáme. Tyto termíny je však možné do metajazyka zavést pouze za pomoci definic. Metajazyk může obsahovat i nedefinované termíny, ale pouze pod podmínkou, že jsou tyto termíny obsaženy již v objektovém jazyku a jsou tak (nutnou) částí metajazyka. Druhou výjimku tvoří termíny „odkazující k formě výrazů objektového jazyka, kterých je použito k tvorbě jmen těchto výrazů“. ⁷⁵ Tarski zdůrazňuje, že je nezbytně nutné, aby byly do jazyka zaváděny sémantické termíny, které by odkazovaly k výrazům objektového jazyka, pouze definicemi. Kdyby tomu tak nebylo, hrozilo by obnovení paradoxu lháře.

Je důležité si uvědomit, že jak objektový jazyk, tak i metajazyk mají relativní smysl. Relativní v tom ohledu, že pokud je zkoumán nějaký objektový jazyk O , pak je toto zkoumání prováděno v metajazyku M^1 , který obsahuje (překlad) O jako svoji součást. Pokud se stane cílem zkoumání metajazyk M^1 , pak bude muset být toto zkoumání provedeno v jiném metajazyku M^2 . Metajazyk M^2 pak bude muset taktéž vyhovovat všem požadavkům kladeným Tarským na metajazyky, a tedy bude muset obsahovat překlad M^1 jako svoji součást. Tím se z M^1 stal ve vztahu k M^2 objektový jazyk. Jestliže by měl být zkoumán podroben M^2 , pak by se stal sám objektovým jazykem pro nový metajazyk M^3 . Toto by mohlo pokračovat do nekonečna. Nemáme zde tedy jediný objektový jazyk a metajazyk, ale rovnou hierarchii metajazyků. Tato hierarchie metajazyků je ale pouze vedlejším produktem této koncepce. Tarski nepostuluje nutnost vytvořit si při zkoumání jazyků jejich hierarchii, a už vůbec ne, že by tato hierarchie měla mít nějaký určitý počet jazyků. Jde mu především o to, že pokud se vyjadřujeme o větách daného objektového jazyka, pak se nacházíme v metajazyku k tomuto jazyku. Kdybychom totiž naše výpovědi o zkoumaném jazyku neformulovali v metajazyku, ale bylo by možné je formulovat přímo ve zkoumaném objektovém jazyku, pak by tento jazyk nevyhovoval Tarského požadavkům na jazyk se specifikovanou strukturou, protože by se jednalo o jazyk sémanticky uzavřený.

⁷⁴ TARSKI, Alfred. Sémantická koncepce pravdy a základy sémantiky. s.148.

⁷⁵ Tamtéž. Tarski zde má na mysli pravděpodobně prostředky potřebné k vytvoření strukturálně-deskriptivních jmen.

Ještě než přistoupím k samotné Tarského definici pravdy, je zapotřebí zmínit se o důležitém principu, který je známý pod jménem Tarského teorém.⁷⁶ Vychází z Teorému I. obsaženého v [CTFL]⁷⁷ a jeho důsledek se dá formulovat následovně: „Za jistých obecných předpokladů je dokázáno, že je nemožné zkonstruovat korektní definici pravdy, pokud jsou použity pouze takové kategorie, které se vyskytují ve zkoumaném jazyce.“⁷⁸ Tedy jinými slovy, podle Tarského teorému je nemožné zkonstruovat adekvátní definici pravdy pro objektový jazyk, pokud používáme pouze prostředky objektového jazyka. Právě toto je důvod, proč musí být adekvátní definice termínu „pravdivý“ provedena v metajazyku, který navíc musí obsahovat sémantické pojmy aplikovatelné na výrazy (především na věty) objektového jazyka.

Než se však dostanu k samotné formulaci definice termínu „pravdivý“, bude třeba vyložit některé doplňující požadavky, kterých je zapotřebí se přidržovat, aby byla výsledná definice adekvátní. Úspěch konstrukce definice pravdy je závislý na tom, jestli je metajazyk, ve kterém bude definice provedena, „bytostně bohatší“ než jazyk, pro který je definice konstruována. Bohužel o podmínce bytostné bohatosti se Tarski příliš podrobně nevyjadřuje. Jako kompenzace tohoto však může posloužit Tarského ukázka, co by měl takový jazyk obsahovat, nacházející se v [SCT]: „Omezíme-li se na jazyky založené na logické teorii typů, pak je jazyk „bytostně bohatší“ než objektový jazyk, jestliže obsahuje proměnné vyšších logických typů, než jsou logické typy proměnných objektového jazyka.“⁷⁹ Metajazyk, který by nevyhovoval této podmínce, by bylo možné rekonstruovat v objektovém jazyku. Tudíž by v takovém metajazyku bylo možné vytvořit paradox plodící větu.

⁷⁶ Nebo také Tarského teorém nedefinovatelnosti, srovnej GÓMEZ-TORRENTE, Mario. The indefinability of truth in the “Wahrheitsbegriff”. *Annals of Pure and Applied Logic* 126. 2004, 27–37.

⁷⁷ TARSKI, Alfred. The Concept of Truth in Formalized Languages. s. 247.

⁷⁸ „Under certain general assumptions, it proves to be impossible to construct a correct definition of truth if only such categories are used, which appear in the language under consideration.“ (TARSKI, Alfred. The Concept of Truth in Formalized Languages. s. 254.)

⁷⁹ TARSKI, Alfred. Sémantická koncepce pravdy a základy sémantiky. s. 140.

4 Konstrukce definice pravdy

Cílem této kapitoly bude pojednat o samotné Tarského definici pravdy. Tarského cílem je použít ke konstrukci definice pojem splňování. Splňování je sémantický termín, což by mohlo představovat pro Tarského definici problém, jelikož metajazyk (ve kterém je konstrukce prováděna) nesmí obsahovat žádné nedefinované sémantické termíny. Pojem splňování je však podle Tarského snadné definovat v nesémantických termínech a lze tak vyhovět předchozímu požadavku. Důležitost toho, aby byla definice pravdy provedena v nesémantických pojmech, shrnuje Milne následovně:

pokud může být pravda — a ostatní sémantické pojmy — definována v nesémantických termínech, potom jsou sémantické pojmy (a) očištěny od náznaku paradoxu dlícího nad nimi, (b) vysvětleny z hlediska pojmů, jejichž význam je jasný a (c) vypadají, že vyhovují požadavkům fyzikalismu⁸⁰

Nyní tedy adekvátnost definice závisí na tom, jak se Tarskému povede definovat termín „pravdivý“ za pomoci pojmu splňování. „Splňování je relace mezi libovolnými předměty a jistými výrazy, které se nazývají „*větné funkce*“.“⁸¹ Tento citát však není úplně přesný, jelikož splňování je nejen relací mezi větnými funkcemi a předměty, ale i mezi větami a předměty. Jelikož věty/větné funkce jsou výrazy jazyka, potom je splňování relací mezi výrazy jazyka a předměty, což odpovídá Tarského definici sémantických termínů z počátku [SCT].⁸²

Větná funkce⁸³ je takový výraz, který má obdobnou formální strukturu, jako má věta. Zatímco věty mohou obsahovat výhradně vázané proměnné, větné funkce mohou obsahovat i volné proměnné. Kirkham k tomuto dodává následující: „Otevřená věta není věta, protože ve skutečnosti nic neříká.“⁸⁴ Aby se větná funkce stala větou (a tudíž získala pravdivostní hodnotu), musí být všechny volné proměnné v dané větné funkci nahrazeny jmény objektů.

Definování větné funkce v sobě nese jistá specifika, kterým musí být podřízena použitá metoda definice. Tarski proto používá metodu rekurzivní definice. Důvod, proč

⁸⁰ „if truth — and other semantic notions — can be defined in non-semantic terms then semantic notions are (a) cleared of the suspicion of paradox that lingers over them, (b) explained in terms of concepts whose meaning is clear, and (c) seem to satisfy the demands of physicalism“ (MILNE, Peter. Tarski on Truth and Its Definition. s. 11.)

⁸¹ TARSKI, Alfred. Sémantická koncepce pravdy a základy sémantiky. s. 150.

⁸² Srovnej tamtéž, s. 141.

⁸³ Kirkham používá místo termínu „větná funkce“ poměrně výstižný termín „otevřená věta“ [open sentence].

⁸⁴ „An open sentence is not a sentence, because it really does not say anything.“ (KIRKHAM, Richard L. *Theories of Truth : A Critical Introduction*. s. 151.)

zvolit právě tuto metodu, je prostý — větných funkcí je nekonečně mnoho a jejich výčet by tak byl nemožný. Rekurzivní definice spočívá v tom, že nejdříve definujeme nejjednodušší větné funkce a posléze určíme pravidla pro jejich spojování.⁸⁵ Pokud máme definovány elementární větné funkce a pravidla pro jejich slučování, nic nám nebrání vytvořit za pomoci těchto pravidel libovolně dlouhou větnou funkci. Obecně lze rekurzivní definici charakterizovat takto: „jedna nebo více klauzulí netautologicky definují nejzákladnější členy definované třídy, jsou následovány jednou nebo více rekurzivními klauzulemi, definujícími, jak jsou další členy třídy vystavěny ze základnějších členů.“⁸⁶ Ovšem, jak poukazuje Kirkham,⁸⁷ i rekurzivní metoda má jedno slabé místo. Rekurzivní definici je totiž možné úspěšně vytvořit jen tehdy, když existuje pouze konečné množství elementárních členů dané třídy. Stejně tak nemůže být rekurzivní definice vytvořena, pokud existuje nekonečné množství pravidel, podle nichž je možné kombinovat elementární členy dané třídy.

Jakmile jsou definovány větné funkce, je možné přejít k definici pojmu splňování. Jak jsem uvedl výše, k tomu, aby se z větné funkce stala věta, musí dojít k nahrazení volných proměnných jmény objektů.

dané předměty splňují danou funkci, jestliže se z této funkce stane pravdivá věta, když v ní nahradíme volné proměnné jmény těchto předmětů.⁸⁸

Tedy pokud máme větnou funkci „x je bílý“, pak je tato funkce splňována např. sněhem, jelikož dosazením jména „sníh“ do zmíněné větné funkce získáme pravdivou větu „sníh je bílý“. Další postup je prakticky stejný, jako v případě definování větné funkce. K definování pojmu splňování je tak opět využita metoda rekurzivní definice. Prvním krokem jsou určeny předměty splňující nejjednodušší větné funkce. Když jsou tyto určeny, pokračuje proces určením podmínek, kterým musejí předměty splňující elementární větné funkce vyhovovat, aby mohly splňovat složené větné funkce.⁸⁹ Problém nastává, když má větná funkce více proměnných, pak totiž přestane být splňování dvojčlennou relací, ale stane se relací troj- a vícečlennou (podle počtu proměnných ve větné funkci). Proto je podle Tarského vhodné upravit definici splňování tak, že splňování je definováno jako

⁸⁵ Srovnej tamtéž, s. 146–150.

⁸⁶ „one or more clauses noncircularly defining the most basic members of the set being defined, followed by one or more recursive clauses defining how other members of the set are built out of more basic members.“ (Tamtéž, s. 148.)

⁸⁷ Srovnej tamtéž, s. 149.

⁸⁸ TARSKI, Alfred. Sémantická koncepce pravdy a základy sémantiky, s. 150.

⁸⁹ Srovnej TARSKI, Alfred. The Concept of Truth in Formalized Languages. s. 190–191.

dvojčlenná relace mezi větnou funkcí a nekonečnou posloupností objektů. Jelikož věty neobsahují žádné proměnné, mohou nastat pouze dva případy — věta je splňována všemi posloupnostmi nebo není splňována žádnými. Definici pravdivé věty lze tedy formulovat následovně: Věta je pravdivá tehdy, když je splňována všemi (nekonečnými) posloupnostmi objektů.⁹⁰

⁹⁰ Srovnej tamtéž, s. 190–198.

5 Adekvátnost pro řešení paradoxů

Na závěr této části hodlám pojednat o tom, jestli Tarského koncepcí vyhovuje požadavkům na adekvátní řešení paradoxů, jak je uvádí Kirkham.⁹¹ Tím si vytvořím jakýsi „most“ pro přechod ke druhé části práce.

1. *Specifičnost* — Řešení musí ukázat, v čem spočívá možnost tvorby paradoxu. Tarski jako příčinu paradoxů uvedl sémantickou uzavřenost jazyků, ve kterých jsou formulovány paradox plodící věty. Proti tomuto bodu je namířena kritika Roberta L. Martina.
2. *Absence ad hoc postulátů* — Nabízené řešení by nemělo obsahovat takové prvky, které byly přijaty pouze na základě toho, že řeší paradoxy. Tarskému je v tomto ohledu vytýkáno zavedení hierarchie metajazyků s pravdivostními predikáty relativizovanými k jazykům, na něž jsou aplikovány. Kritiku tohoto Tarského přístupu podal např. Saul Kripke.
3. *Zamezení přílišným restrikcím* — Řešení paradoxu by mělo být takové, aby nebyly na jeho úkor odstraněny i neproblematické věty. Důsledkem odmítnutí sémanticky uzavřených jazyků však je, že jsou z objektového jazyka odstraněny veškeré autoreferenční věty. Kritiku tohoto bodu lze nalézt u R. L. Martina.
4. *Kompletnost* — Nabízené řešení by mělo být aplikovatelné na všechny varianty paradoxu. Jestli Tarski tomuto kritériu vyhovuje, se pokusím určit za pomoci prací dr. Jiřího Raclavského věnovaných paradoxům lháře.
5. *Soulad s intuicemi* — Řešení paradoxu by se nemělo protivit našim intuicím o významu termínu „pravdivý“. Kritikové Tarskému vytýkají, že hierarchie metajazyků a s nimi i hierarchie pravdivostních predikátů neodpovídají běžnému úzu. Do této skupiny kritiků by se dal zařadit např. Jonathan Harrison, podle kterého není v souladu s našimi intuicemi aplikovat termín „pravdivý“ na věty, a Saul Kripke se svou kritikou zmnožení pravdivostních predikátů v Tarského koncepci.⁹²

Tímto je výkladová část této práce vyčerpána a nezbyvá, než přejít k její kritické části.

⁹¹ KIRKHAM, Richard L. *Theories of Truth: A Critical Introduction*. s. 273.

⁹² Volně podle KIRKHAM, Richard L. *Theories of Truth: A Critical Introduction*. s. 273.

Kritika Tarského sémantické koncepce pravdy

6 Problémy hierarchie metajazyků

6.1 Mstivý lhář a nedostatečnost hierarchií metajazyků

V první části jsem se krátce zmínil o paradoxu lháře a o Tarského přístupu k tomuto paradoxu. Nyní se k němu hodlám vrátit. Tedy lépe řečeno vrátím se k jedné jeho variantě, která bývá používána jako argument proti části Tarského koncepce. Jak jsem již naznačil dříve, paradox lháře je mezi logiky velmi oblíbený a s postupem času se objevilo velké množství jeho variant.⁹³ Některé jsou „pouze“ elaborací klasického lháře, jiné vznikly se záměrem napadnout různé filosofické a logické koncepce — pro účely této kapitoly se budu zabývat pouze druhou uvedenou skupinou.

Řeč nyní bude o takzvaném mstivém lháři [revenge liar], který patří spolu se zesíleným lhářem [strengthened liar] mezi tzv. silnější lháře [stronger liar].⁹⁴ Oproti zesílenému lháři není mstivý lhář zaměřen pouze na klíčový predikát řešení klasického lhářského paradoxu, ale je zaměřen proti principu daného řešení jako celku. Principem, na který je mstivý lhář zaměřen, je existence nedostatečného počtu hierarchií metajazyků. Existují přinejmenším dva eventuální způsoby, jak mstivého lháře zformulovat. Jak se pokusím ukázat dále, obě tyto formulace jsou ve své podstatě zakotveny ve stejném (a dle mého názoru chybném) předpokladu a ani jedna nepředstavuje žádné nebezpečí jak pro koncepci metajazyků, tak pro Tarského koncepci jako celek.

⁹³ Různé varianty paradoxu lháře, spolu s návrhem jejich řešení založeném na systému transparentní intenzionální logiky navržené Pavlem Tichým lze nalézt v:

RACLAVSKÝ, Jiří. Paradox lháře a jeho řešení [online]. 2007 [cit. 2010-02-12]. Dostupné z:

<http://www.phil.muni.cz/~raclavsk/texty/paradox_lhare.pdf>.

potažmo v

RACLAVSKÝ, Jiří. Lhářský paradox, význam a pravdivost [online]. 2009 [cit. 2010-02-12]. Dostupné z:

<http://www.phil.muni.cz/~raclavsk/texty/lharske_paradoxy.pdf>.

⁹⁴ „Zesíleným lhářem se rozumí taková věta, která vrací do hry lhářský paradox (ve smyslu paradoxní inference) tím, že ona věta obsahuje predikát, který je klíčovým predikátem příslušného řešení.“ (RACLAVSKÝ, Jiří. Lhářský paradox, význam a pravdivost. s. 20.)

Zesílený lhář je tedy hrozbou především pro systémy vícehodnotové logiky. Jeho příkladem může být:

SL: Věta (SL) je významuprostá v L^k nebo je nepravdivá v L^k . (Kde L^k je jazyk, ve kterém je věta pronesena).

Jelikož Tarski rozhodně nepatřil mezi zastánce troj- či vícehodnotových logik a v jeho koncepci se nic takového jako třetí pravdivostní hodnota neobjevuje, je jasné, že tento paradox Tarského systém nijak neohrožuje. Ovšem Kripkeho teorie pravdy rozebíraná na konci této části práce už tímto paradoxem napadnutelná je.

První varianta mstivého lháře má, dle mého názoru, nejlépe viditelné nedostatky a je snáze napadnutelná. Ostatně posuďte sami:

(RL₁) Věta (RL₁) není pravdivá v žádném z jazyků J¹–Jⁿ.⁹⁵

Na první pohled je jasné, že tato varianta je zaměřena na (meta)jazyky řádů 1–n. Princip, kterým má tento paradox zpochybnit Tarského koncepci je následující:

- (1) Máme dány pouze jazyky řádů 1–n, tedy J¹–Jⁿ
- (2) Věta (RL₁) tedy musí být formulována v jednom z jazyků J¹–Jⁿ
- (3) Věta (RL₁) o sobě vypovídá, že není pravdivá ani v jednom z jazyků J¹–Jⁿ
- (4) Avšak k tomu, aby mohla věta (RL₁) vypovídat o jazycích řádu 1–n, by sama musela být v jazyku vyššího řádu, tedy alespoň Jⁿ⁺¹
- (5) Z výše uvedeného však dostáváme spor, jelikož (RL₁) by měla být zároveň v některém z jazyků J¹–Jⁿ a zároveň v jazyku Jⁿ⁺¹
- (6) Hierarchie metajazyků vede k paradoxu a je tudíž nedostačující

To vypadá jako platná námitka, protože věta (RL₁), vyjadřující se o větách řádu 1–n, sama vypadá jako věta řádu n+1. Pokud však můžeme použít pouze jazyky n řádů, vzniká nám problém.

Druhá formulace tohoto paradoxu vypadá nebezpečněji, ačkoli je její znění velmi podobné:

(RL₂) Věta (RL₂) není pravdivá na žádné úrovni (dané) hierarchie jazyků.

Oproti předchozí formulaci se v této neobjevuje odkaz na pevně daný počet řádů jazyků. Ovšem paradox plodící schéma je zde prakticky stejné jako v předchozím případě. V této formulaci nenalezneme žádný explicitní odkaz na počet hierarchií (meta)jazyků, je tedy zapotřebí počítat s „nejhorší“ variantou, tedy s jejich nekonečným množstvím. Problém by pak byl prakticky stejný, jako v předchozím případě, akorát by zde bylo zapotřebí operovat s jazyky nekonečného množství řádů. Výsledek by však byl stejný jako v předchozí variantě. I kdyby byla hierarchie jazyků nekonečná, stejně by se v případě věty (RL₂) zdálo, že ani nekonečná hierarchie nestačí. Mstivý lhář tak napadá hierarchie jazyků tvrzením, že ať je jejich množství jakékoliv, nikdy není dostatečné. Nyní se pokusím ukázat, proč není tato námitka proti Tarského koncepci adekvátní.

Již dříve jsem uvedl, že mstivý lhář je založen na předpokladu, že disponujeme určitým množstvím metajazyků. Jenomže s tímto předpokladem mstivý lhář coby námitka

⁹⁵ Volně podle RACLAVSKÝ, Jiří. Lhářský paradox, význam a pravdivost. s. 21.

proti Tarskému padá. Problémem je právě daný počet metajazyků. Tarski nikde netvrdí, že by mohl existovat pouze jistý počet metajazyků a také nikde neuvádí, že v tom kterém případě máme k dispozici konkrétní počet metajazyků. Z jeho koncepce je sice možné odvodit existenci hierarchie metajazyků (která může být nekonečná), ale to ještě samo o sobě neospravedlňuje napadení jeho koncepce mstivým lhářem. K pochopení, proč tomu tak je, se stačí znovu podívat na závěr kapitoly 3.4, věnované distinkci mezi objektovým jazykem a metajazykem. Jak jsem tam totiž uvedl, pro Tarského má distinkce mezi objektovým jazykem a metajazykem relativní smysl. Pokud mluvíme o libovolném jazyku, vždy o něm musíme mluvit v metajazyku, který bude dostatečně bohatý, aby v něm mohly být formulovány výpovědi o všech větách zkoumaného jazyka. K tomu, aby bylo možné zformulovat smysluplné a konzistentní výpovědi o větách nějakého jazyka, je zapotřebí nacházet se v metajazyku, který bude obsahovat jak jména a překlady všech výrazů zkoumaného jazyka, tak i sémantické termíny aplikovatelné na výrazy tohoto jazyka. Jestliže chceme mluvit o metajazyku M^2 , pak musíme mluvit v metajazyku M^3 (který je esenciálně bohatší než M^2). Chceme-li mluvit o M^3 , musíme vystoupit do (opět esenciálně bohatšího) M^4 a tak dále (klidně až do nekonečna). Vzniká tak sice hierarchie metajazyků, ale to není základem Tarského koncepce: „Jenže podstatou hierarchických řešení není určité (omezené) číslo hierarchií, ale princip komentujících „metapozicí“.“⁹⁶

Jaká je tedy odpověď na otázku, proč je mstivý lhář v tomto případě neadekvátní? Mstivého lháře je nepřipustné aplikovat na Tarského systém, jelikož princip, který napadá — postulování určitého počtu jazyků —, není klíčovým principem tohoto systému. A co víc, postulování určitého počtu metajazyků nejen, že není klíčovým principem Tarského koncepce, ale vůbec se v ní nevyskytuje. Klíčovým principem Tarského koncepce je nutnost vystoupit do metajazyka a definici pravdy pro objektový jazyk zkonstruovat až v metajazyku. Tím však není řečeno, že musíme použít určité množství jazyků, jak vyžaduje mstivý lhář.

6.2 Nerozhodnutelnost zařazení k metajazyku

Jak jsem již uvedl ve výkladové části této práce, mstivý lhář není jedinou variantou paradoxu lháře, kterou je argumentováno proti Tarského koncepci. Nyní se zaměřím na dva vzájemně velmi podobné paradoxy, jejichž kritice může být Tarského systém

⁹⁶ Tamtéž.

vystaven. Stejně jako v předchozím případě zastávám názor, že oba tyto paradoxy nepředstavují žádné nebezpečí pro Tarského koncepci, což se pokusím i zdůvodnit.

Stejně jako v předchozí kapitole jsou i tyto dva paradoxy variantami paradoxu lháře. Zatímco asi nejslavnější formulací paradoxu lháře je jeho autoreferenční nekvantifikační varianta, v této kapitole se budu zabývat variantami heteroreferenčními.⁹⁷ Začnu heteroreferenční nekvantifikační variantou paradoxu lháře, ve znění:

JP₁: Věta JP₂ je nepravdivá.

JP₂: Věta JP₁ je pravdivá.⁹⁸

Na rozdíl od předchozí kapitoly jsou obě varianty heteroreferenčního paradoxu lháře velmi podobné a jak jejich důsledky, tak i jejich nedostatky jsou prakticky stejné. Proto nyní uvedu rovnou heteroreferenční kvantifikační variantu a hodnocení obou paradoxů pak uvedu společně.

Platón: Vše, co řekl Sókratés, je pravdivé.

Sókratés: Vše, co řekl Platón, je nepravdivé.⁹⁹

Tyto dvě obměny znázorňují, že paradoxní situace může vzniknout i jinak, než za pomoci autoreference, jak je tomu např. u klasické varianty paradoxu lháře. Zde máme v obou případech dvě věty, přičemž ani jedna není autoreferenční a tedy sama o sobě nebezpečná. Problém ovšem nastává, když se tyto věty uvažují dohromady. Díky jejich spojení dojde k zacyklení — jedna věta odkazuje k větě druhé a druhá věta zase odkazuje k větě první. Schéma plodící paradox je pro nekvantifikační variantu následující:

- (1) Předpokládáme, že věta (JP₁) je pravdivá
- (2) Jestliže je věta (JP₁) pravdivá, pak je pravdivé to, co říká
- (3) Věta (JP₁) říká, že věta (JP₂) je nepravdivá
- (4) Věta (JP₂) je tedy nepravdivá
- (5) Jestliže je (JP₂) nepravdivá, pak to, co říká, není pravda

⁹⁷ Robert L. Martin používá k označení variant paradoxu lháře, které já označuji jako heteroreferenční, termín nepřímá autoreference [indirect self-reference]. Viz MARTIN, Robert L. Toward a Solution to the Liar Paradox. *The Philosophical Review*. July 1967, vol. 76, no. 3, s. 281.

⁹⁸ RACLAVSKÝ, Jiří. Lhářský paradox, význam a pravdivost. s. 23.

Tato varianta je též známá jako Jourdainův paradox (po britském logikovi Philipu E. B. Jourdainovi). Heteroreferenční nekvantifikační variantu paradoxu lháře je možné zkonstruovat i následujícími dvěma způsoby:

Dvojice vět, z nichž první je „Následující věta je nepravdivá“ a druhá je „Předcházející věta je pravdivá“.

Nebo máme kartičku, na jejíž jedné straně je nápis „Věta na druhé straně kartičky je nepravdivá“ a na druhé straně je nápis „Věta na druhé straně kartičky je pravdivá“.

⁹⁹ Podle RACLAVSKÝ, Jiří. Lhářský paradox, význam a pravdivost. s. 24.

Této variantě se říká taktéž buridanovská insolublie. Své jméno tento paradox získal po středověkém knězi a filosofovi Jeanu Buridanovi, který se zabýval mimo jiné i paradoxy.

- (6) Věta (JP_2) říká, že věta (JP_1) je pravdivá
- (7) Podle (5) a (6) je tedy věta (JP_1) nepravdivá
- (8) (JP_1) je tedy pravdivá a nepravdivá zároveň

Je samozřejmě možné začít předpokladem pravdivosti věty JP_2 , nebo i předpokladem nepravdivosti vět JP_1 nebo JP_2 — konstrukci odpovídajících schémat ponechám na čtenáři. Výsledek však bude vždy stejný – jedna z těchto vět bude pravdivá i nepravdivá zároveň, čímž vznikne paradoxní situace. Samozřejmě, že prakticky stejné schéma je možné vytvořit i pro buridanovskou insolubilitu — tento krok opět ponechám na čtenáři.¹⁰⁰

Zatímco u mstivého lháře jde o napadení hierarchie metajazyků z důvodu jejich údajně nedostatečného počtu, heteroreferenční lhářské paradoxy mají ukázat, že v některých případech je nemožné určit, ke kterému metajazyku (respektive k jazyku kterého řádu) patří ta která věta v jejich zadání. Pro pochopení této námítky je vhodné ilustrovat ji dalším schématem:

- I. Věta (JP_1) zmiňuje větu (JP_2)
- II. Aby se mohla nějaká věta zmiňovat o jiné větě, musí být formulována v (meta)jazyku vyššího řádu
- III. Předpokládáme, že (JP_2) je v jazyku řádu k (J^k)
- IV. Aby mohla (JP_1) zmiňovat (JP_2), musí být formulována alespoň v jazyku řádu $k+1$ (J^{k+1})
- V. (JP_1) je tedy v J^{k+1} a (JP_2) v J^k
- VI. (JP_2) zmiňuje (JP_1)
- VII. (JP_2) tedy musí být podle (II) formulována v jazyku vyššího řádu, než (JP_1), tedy alespoň v J^{k+2}
- VIII. Z (V) a (VII) vzniká spor, jelikož (JP_2) by měla být formulována v J^k a J^{k+2} zároveň

Na základě tohoto schématu je tedy údajně oprávněné tvrdit, že v Tarského systému není možné rozhodnout, ve kterém jazyku jsou zformulovány věty heteroreferenčního paradoxu. Ovšem to by platilo pouze, pokud by bylo výše uvedené schéma adekvátním rozbohem heteroreferenčního lháře. Ve zbytku této kapitoly se pokusím ukázat, že tomu tak není a Tarského systém proto není heteroreferenčními variantami paradoxu lháře ohrožen. Stejně jako v předchozí kapitole toto řešení pochází od dr. Raclavského.¹⁰¹

Nejdříve je důležité uvědomit si, že výše uvedené schéma počítá pouze s jedním možným způsobem, jak lze údajně zkonstruovat heteroreferenční variantu paradoxu lháře, kde východiskem je předpoklad, že věta (JP_1) je formulována v jazyku J^{k+1} a věta (JP_2) je

¹⁰⁰ Srovnej tamtéž, s. 24–25.

¹⁰¹ Srovnej tamtéž, s. 23–24.

v jazyku J^k — tuto variantu označím jako (α) . Je samozřejmě možné tyto předpoklady prohodit a začínat tak s větou (JP_2) formulovanou v jazyku J^{k+1} a naopak s větou (JP_1) v jazyku J^k (schéma by i přes tuto změnu zůstalo prakticky stejné) — tuto variantu označím jako (β) . Nabízí se také možnost vyjít z předpokladu, že jak věta (JP_1) , tak i věta (JP_2) jsou formulovány v jazyku stejné úrovně, např. J^k — varianta (γ) —, ale jak ukážu později, toto je nejproblematictější možnost a její neadekvátnost se projeví dříve než u předchozích případů.

Jelikož důkazy neadekvátnosti (α) a (β) se od sebe liší pouze prohozením jednotlivých vět, budu se zde zabývat pouze důkazem neadekvátnosti (α) a důkaz pro (β) ponechám na čtenáři. Začátek důkazu je stejný, jako ve výše uvedeném schématu — věta (JP_2) je formulována v jazyku J^k a věta (JP_1) , která ji zmiňuje, je tudíž v jazyku J^{k+1} . Dosud zde není žádný problém — ten ale vyvstává hned v dalším kroku (VI). O větě (JP_1) jsme oprávněni tvrdit, že zmiňuje větu (JP_2) , jelikož je formulována v jazyku vyššího řádu. Je totiž formulována v jazyku J^{k+1} , který obsahuje prostředky umožňující pojmenovat a tedy i zmiňovat všechny věty jazyka nižšího řádu, tedy i J^k . Věta (JP_2) , která je sama formulována v J^k obsahuje odkaz k (JP_1) — to však v žádném případě nesmí. Žádný jazyk nesmí podle Tarského obsahovat kromě jmen pro vlastní věty i jména vět jazyků vyššího řádu. To je ale v případě věty (JP_2) zcela zjevně porušeno. Větu (JP_2) je tedy nutné považovat za nesmyslnou a to bez ohledu na to, co o ní říká věta (JP_1) . Tímto ovšem mizí problém se zacyklením a žádný paradox vlastně nevzniká.

V případě (γ) je situace ještě jednodušeji vyřešitelná. Jelikož obě věty jsou v tomto případě formulovány v jazyku J^k , pak ani jedna z nich není smysluplná. Jelikož kdyby tyto věty mohly odkazovat k větám stejného jazyka, ve kterém jsou zformulovány, pak by byly samy zformulovány v sémanticky uzavřeném jazyku. A to Tarski rozhodně nepřipouští.

Doufám, že tímto jsem dostatečně vysvětlil, proč nepředstavuje heteroreferenční nekvantifikační varianta paradoxu lháře žádné nebezpečí pro Tarského sémantickou koncepci pravdy. Jelikož důkaz neškodnosti heteroreferenční kvantifikační varianty by vypadal prakticky shodně, ponechám jeho studium na čtenáři.¹⁰²

¹⁰² Srovnej tamtéž, s. 24–25.

7 Záchrana autoreferenčních vět

Jak bylo uvedeno již v první části této práce, Tarski se zabýval problémem definování pravdy pouze pro jazyky se specifikovanou strukturou a odmítl všechny jazyky sémanticky uzavřené. Důsledkem toho je, že Tarského definice jsou aplikovatelné pouze pro jazyky, které neobsahují žádné autoreferenční výrazy — jelikož prostředky pro odkazování k výrazům jazyka J^k mohou být v Tarského koncepci pouze v jazyku vyššího řádu J^{k+1} . Podle některých je toto omezení příliš restriktivní a Tarski tak odstraňuje i množství vět, které jsou neproblematické. To je podle kritiků příliš velká cena za odstranění paradoxů. V této kapitole rozeberu přístup reprezentanta „populárních“ námitek, Roberta L. Martina, jehož cílem bylo ukázat, že problémy s autoreferenčními větami lze vyřešit, aniž bychom použili hierarchie jazyků a odstranili tak z jazyka autoreferenční výrazy.

Námítka v této kapitole projednávaná má své kořeny v Tarského definici sémanticky uzavřených jazyků. Kritéria, kterým sémanticky uzavřený jazyk vyhovuje, lze nalézt v kapitole 3.2 této práce. Klíčovou roli zde budou hrát kritéria, která jsem tam uvedl pod bodem (a):

- i. Sémanticky uzavřený jazyk obsahuje kromě vlastních výrazů také jména těchto výrazů (nebo jiný způsob umožňující mluvit o větách tohoto jazyka)
- ii. sémantické termíny odkazující k větám tohoto jazyka

Bod (i) nám o sémanticky uzavřených jazycích říká to, že každý takový jazyk v sobě obsahuje prostředky, které mu umožňují mluvit o vlastních větách. Tedy jinými slovy je takový jazyk natolik bohatý, že je v něm možné formulovat věty odkazující k jiným větám tohoto jazyka nebo k nim samým. Jedná se zde tedy o tzv. autoreferenci — jev, jehož vzniku se Tarski brání zavedením pravidla, že prostředky umožňující zmiňovat prvky daného jazyka spolu se sémantickými termíny aplikovatelnými na věty tohoto jazyka, mohou být pouze v jazyku vyššího řádu.

Příkladem nebezpečnosti autoreferenčních vět může být mnou již zmíněný paradox lháře v následujícím znění:

(L) Věta (L) je nepravdivá.

Tarski spatřoval nebezpečí pro adekvátní teorii pravdy jak v bodu (i), tak i v bodu (ii). Pokud chceme zabránit tvorbě paradoxů v daném jazyku L, pak podle něho musíme z tohoto jazyka odstranit všechny jeho sémantické pojmy a prvky umožňující pojmenovávání vlastních výrazů (a přesunout je do metajazyka k tomuto jazyku). Pokud

ovšem z jazyka odstraníme veškeré prostředky umožňující jeho větám zmiňovat samy sebe, pak z tohoto jazyka zmizí veškeré autoreferenční výrazy.

Martinova námitka je založena na tom, že autoreferenční výrazy mohou být jak „dobré“, tak i „špatné“. „Špatnou“ autoreferencí jsou myšleny případy, kdy dochází ke vzniku paradoxů. Tarski založil svoje odmítnutí autoreference právě na těchto případech. Je zde ale ještě druhá skupina autoreferenčních výrazů, která obsahuje věty, které „mohou být mluvčím jazyka snadno pochopeny a mohou být posouzeny logicky neproblematickým způsobem jako jasně pravdivé (nebo nepravdivé)“.¹⁰³ Příkladem „dobré“ autoreference může být např. věta „Tato věta je v češtině“ — tato věta vypovídá sama o sobě, a přesto není zdrojem paradoxů, jelikož je neproblematické určit její pravdivostní hodnotu.

Martin se snaží nabídnout řešení, které by v daném jazyku zabraňovalo tvorbě paradoxů, ale zároveň nebylo tak restriktivní jako Tarského řešení. Navíc se jeho řešení obejde bez toho, aby se sémantické termíny a prostředky pro pojmenovávání výrazů zkoumaného jazyka musely přesouvat do metajazyka. Svoji argumentaci Martin zahajuje poukázáním na jisté rozdíly mezi variantami paradoxu lháře. Uvažme následující možnosti:

1. „Tato věta je nepravdivá.“
2.
 - a. „Následující věta je pravdivá.“
 - b. „Předchozí věta je nepravdivá.“

Paradox plodící větu (1) Martin označuje jako „crux“ verzi paradoxu lháře. Varianta (2) není klasickou autoreferenční variantou, jelikož obsahuje dvě navzájem k sobě odkazující věty. Martin používá pro vztah mezi větami v této variantě termín nepřímá autoreference.¹⁰⁴ Mezi „crux“ verzí a heteroreferenční variantou spatřuje Martin jeden důležitý rozdíl. Zatímco paradox plodící věta varianty (1) k sobě odkazuje naprosto vždy, věty (2a) a (2b) musejí být ve správném kontextu, aby odkazovaly jedna k druhé. Tedy zatímco u věty (1) je jedno, kde bude umístěna, u vět (2a) a (2b) záleží na tom, aby byly umístěny bezprostředně po sobě (i když bez ohledu na jejich pořadí). Z toho lze vyvodit, že zatímco (1) je paradoxní vždy, vznik paradoxu z (2a) a (2b) je závislý na kontextu, ve kterém se věty vyskytují. Kdyby byla mezi věty obsažené ve (2a) a (2b) vložena nějaká další věta, např. „Venku zpívají ptáci“, pak by nedošlo ke vzniku paradoxní situace, jelikož

¹⁰³ „can be readily understood by a speaker of the language, and be judged, in a logically unproblematic way, to be straightforwardly true (or false).“ (MARTIN, Robert L. *Toward a Solution to the Liar Paradox*. s. 279.)

¹⁰⁴ Já jsem doposud pro podobné varianty paradoxu lháře používal termín heteroreferenční.

pravdivostní hodnoty vět (2a) a (2b) by byly závislé na pravdivostní hodnotě věty mezi ně vložené. Martin k tomuto dodává: „Kvůli tomuto rozdílu mezi druhy verzí Lháře budeme rozlišovat mezi „sémantickou nekorektností“ (aplikovatelnou na větné typy) a „tokenovou zvláštností“ (aplikovatelnou na větné tokeny)¹⁰⁵.“¹⁰⁶ V případě „cruX“ varianty paradoxu lháře jde tedy o sémantickou nekorektnost, kdežto u heteroreferenční varianty jde o tokenovou zvláštnost.

Pokud má být věta považována za smysluplnou a pokud jí má být přiřazena pravdivostní hodnota, pak musí taková věta splňovat dva požadavky. Prvním z nich je, aby byla v souladu s pravidly gramatiky — taková věta musí být formulována s ohledem na pravidla syntaxe daného jazyka. Podle tohoto pravidla by byly jako nesmyslné vyloučeny věty jako např. „Paruce v okenice bambusová zpívala měsíční“. Toto pravidlo samo o sobě však nestačí. Je totiž možné formulovat gramaticky korektní větu, která bude i přesto nesmyslná. Takovou větou může být např. „bambusová okenice zpívala v měsíční paruce“. Martin k tomuto dodává, že kromě gramatiky je třeba brát v úvahu i sémantiku.¹⁰⁷

Na věty, které jsou sémanticky nekorektní, není podle Martina možné vztahovat princip bivalence. Pokud bychom jeho dodržování vyžadovali, vedlo by to v případě těchto vět k paradoxním situacím. Martinovou modelovou větou je „Jedenáctka je fialová“. Pokud bychom o ní řekli, že je nepravdivá, pak bychom se dostali do situace, že bychom museli připsat pravdivost některé alternativní větě, jako je „Jedenáctka je žlutá“ nebo „Jedenáctka je červená“.¹⁰⁸ Právě proto by se takovýmto větám neměla připisovat žádná pravdivostní hodnota. Dalším Martinovým krokem je návrh procedury, jejímž cílem by bylo rozpoznat veškeré deviantní věty. Rozebírání této metody by však bylo nad rámec této práce.¹⁰⁹

Martin tak na rozdíl od Tarského neodstraňuje z jazyka veškeré autoreferenční věty, ale pouze ty, které jsou deviantní. Jelikož paradox plodící věty, jako je např. „Tato věta je nepravdivá“, nemohou úspěšně projít jeho rozpoznávací procedurou, není na ně aplikován princip bivalence a není jim tak přiřazena pravdivostní hodnota. Neproblematické (nejenom) autoreferenční věty však touto procedurou projdou a může jim být připsána

¹⁰⁵ Protože paradox v těchto případech vždy plodí pouze tokeny vět, tedy jednotlivé nápisy, potažmo vyřčení. Jedině tokeny vět — jakými jsou věty (2a) a (2b) — jsou totiž vždy napsány nebo vysloveny v nějakém kontextu.

¹⁰⁶ „It is because of this difference between kinds of versions of the Liar that we will distinguish between “semantical incorrectness” (applicable to sentence types) and “token oddity” (applicable to sentence tokens).“ (Tamtéž, s. 282–283.)

¹⁰⁷ Srovnej tamtéž, s. 286.

¹⁰⁸ Tamtéž, s. 289.

¹⁰⁹ Metoda je popsána v MARTIN, Robert L. *Toward a Solution to the Liar Paradox*. s. 291–300.

pravdivostní hodnota. Martin se tak zbavuje problémů s deviantními větami, aniž by se, jako Tarski, zbavoval všech autoreferenčních vět a musel kvůli tomu provést vzestup do metajazyka.

8 Nekonzistence hovorových jazyků

Tato kapitola bude věnována kritice pocházející od Hanse Herzbergera, podle kterého nemůžeme o žádném jazyku vypovědět, že je nekonzistentní, aniž by tato naše výpověď byla kontradikcí. Svoji argumentaci zahajuje Herzberger konstatováním, že pokud má být nějaký jazyk nekonzistentní, pak jediné proto, že je nekonzistentní množina analytických vět v tomto jazyku obsažených. To je ale podle Herzbergera nepřijatelné, což se pokouší dokázat argumentem založeným na *reductio ad absurdum* — tedy tak, že z přijatých premis plynou nepřijatelné závěry.¹¹⁰ Jeho argument má následující schéma:

- 1) Daný jazyk L není logicky konzistentní kvůli své třídě analytických vět
- 2) L obsahuje neprázdnou třídu vět A takovou, že
 - a) každá věta patřící do A je analytická a
 - b) třída A je logicky nekonzistentní třída vět
- 3) Každá věta patřící do třídy A je pravdivá
- 4) Alespoň jedna věta patřící do A musí být nepravdivá
- 5) V jazyku L existuje alespoň jedna věta, která je pravdivá a nepravdivá zároveň¹¹¹

Rozbor tohoto argumentu začnu až u bodu (2). Pokud by byla množina analytických vět zkoumaného jazyka prázdnou množinou, pak by podle Herzbergera nemělo smysl nadále se zabývat zkoumáním konzistence tohoto jazyka. Bod (2b) je přímým důsledkem první premisy, podle které musí být třída analytických vět nekonzistentní. Toho je možné dosáhnout pouze tak, že bude daná třída obsahovat alespoň jednu větu, která bude nepravdivá. Bod (3) je jistým rozvedením bodu (2a), jelikož pokud je A třídou analytických vět, pak z definice pojmu „analytický“ musejí být všechny věty v této třídě obsažené pravdivé. Oproti tomu bod (4) vychází z premis (1) a (2b), podle kterých musí být třída A nekonzistentní a tudíž musí obsahovat alespoň jednu pravdivou větu. Z tohoto ale plyne, že třída A (a tedy i jazyk L) obsahuje alespoň jednu větu, která by byla pravdivá i nepravdivá zároveň. To je ale zjevný rozpor. Zkoumaný jazyk L tak nemůže být nekonzistentní právě proto, že jeho třída analytických vět A může obsahovat pouze analytické, a tedy pravdivé věty.

Další argumentace je založena na zkoumání pravdivostních podmínek. Podle Herzbergera je totiž myšlenka nekonzistence jazyků založena na představě, že „pravdivostní podmínky jazyka by snad mohly být takové, že, za určitých podmínek, by

¹¹⁰ *Reductio ad absurdum* bylo využito taktéž ve výkladech o paradoxu lháře v této práci.

¹¹¹ Schéma volně podle HERZBERGER, Hans G. The Truth-Conditional Consistency of Natural Languages. *The Journal of Philosophy*. February 1967, vol. 64, no. 2, s. 29.

byly logicky nekonzistentní třídy vět učiněny pravdivými“.¹¹² Tedy jinými slovy, přesvědčení některých o nekonzistenci jazyků je založeno na představě, že v určitých situacích existují podmínky pro to, aby byly za pravdivé považovány dvojice vzájemně kontradiktorických vět. Vezme-li se toto v potaz, je prý možné vytvořit podmínky k tomu, aby byl jazyk nekonzistentní na základě svých pravdivostních podmínek. Tyto podmínky by spočívaly v tom, že by nastala situace, mající za následek, že by se pravdivými staly všechny věty, které by patřily do třídy logicky nekonzistentních vět. Herzberger však vyvrací i tuto možnost vytvoření nekonzistentního jazyka.

- I. Máme daný jazyk L , který obsahuje logicky nekonzistentní třídu vět Q
- II. Je dána podmínka w taková, že pokud w nastane, učiní pravdivými všechny věty Q
- III. Jelikož Q je logicky nekonzistentní třída vět, je nemožné, aby byly všechny věty v ní obsažené pravdivé
- IV. Jelikož je nemožné, aby byly všechny věty Q pravdivé, je nemožné, aby nastala podmínka w ¹¹³

Jelikož neexistují podmínky, za kterých by mohly být pravdivé všechny věty patřící do logicky nekonzistentní třídy vět, není možné o žádném jazyku prohlásit, že by byl nekonzistentní na základě pravdivostních podmínek. Tento druhý argument je podle Herzbergera generalizací prvního argumentu. Zatímco v tomto argumentu bylo prokázáno, že neexistují podmínky, za nichž by mohly nabýt pravdivost vzájemně kontradiktorické věty, první byl postaven na pojmu analytické věty, která je pravdivá za každých podmínek (a tedy věty, která nemůže být za žádných podmínek nepravdivá). Proto podle Herzbergera „každý jazyk, který je konzistentní na základě pravdivostních podmínek, bude tímto konzistentní i svými analytickými větami“.¹¹⁴

Z výše uvedeného je zřejmé, jak chápe nekonzistentní jazyky Herzberger. Odpověď na otázku, zda stejně chápal nekonzistentní jazyky i Tarski, je nejen dle mého názoru záporná. Jak ostatně uvádí i Soames,¹¹⁵ pro Tarského nebyl jazyk nekonzistentní kvůli nekonzistenci třídy svých analytických vět, neboť Tarski nepovažoval ekvivalence tvaru (T) za analytické věty daného jazyka. Jeho požadavkem bylo, aby ekvivalence tvaru (T) vzniklé dosazením vět do T-schématu byly pravdivé — požadavek, aby tyto věty byly

¹¹² „the truth conditions of a language might conceivably be such that, under certain conditions, logically inconsistent sets of sentences would be rendered true.“ (Tamtéž, s. 31.)

¹¹³ Volně podle HERZBERGER, Hans G. *The Truth-Conditional Consistency of Natural Languages*. s. 32–33.

¹¹⁴ „any language that is truth-conditionally consistent will therefore be consistent in its analytic sentences.“ (Tamtéž, s. 33.)

¹¹⁵ Srovnej SOAMES, Scott. *Understanding Truth*. s. 62–64.

analytické, nevznáší a ani není z jeho požadavků odvoditelný.¹¹⁶ Třída analytických vět se tak i v případě dosazení paradox plodící lhářské věty do T-schématu nestává nekonzistentní. Je tak možné říci, že Herzbergerova kritika se poněkud mívá účinkem, jelikož u Tarského není hovorový jazyk nekonzistentní proto, že by byla nekonzistentní třída jeho analytických vět.

¹¹⁶ Srovnej TARSKI, Alfred. The Concept of Truth in Formalized Languages. s. 165.

9 Aplikovatelnost predikátu „pravdivý“ na věty

V první části jsem se sice již problematice aplikace pravdivostního predikátu na věty věnoval, ale nezabýval jsem se přímo kritikou tohoto Tarského rozhodnutí. To hodlám učinit až na tomto místě. Předložím zde argumenty, které proti možnosti aplikovatelnosti predikátu „pravdivý“ na věty předložil Jonathan Harrison.

Postoj zastávaný Harrisonem lze asi nejlépe sumarizovat tak, jak to udělal jeho kritik Daniel Boisvert: „Podle Harrisona musí být Tarského teorie chybná, protože (a) pravda nemůže být vlastností vět; (b) kdyby pravda mohla být vlastností vět, pak by T-věty musely být nutné pravdy, což nejsou; (c) nejenom, že T-věty nejsou nutné pravdy, ony mohou být i nepravdivé.“¹¹⁷

Harrisonovo stanovisko je, že Tarski a všichni zastánci T-schématu se dopouštějí chyby, když považují za pravděnce věty, jelikož věty být pravděnci nemohou. Celý svůj článek tak věnuje tomu, aby dokázal, že věty opravdu nejsou vhodnými kandidáty na roli pravděnců. Na otázku, jaká entita by měla mít roli pravděnců, odpovídá Harrison poměrně jasně — propozice. O propozicích prohlašuje Harrison, že je chápe jako „věci“, které jsou, nebo mohou [...] být tvrzeny a které, kdyby byly tvrzeny, by byly pravdivé bez ohledu na slova, v nichž byly formulovány“.¹¹⁸ Pročež je podle něho neadekvátní mluvit o tom, že věta je pravdivá nebo nepravdivá. Úkolem vět je vyjadřovat propozice.

Pro pochopení další Harrisonovy argumentace je důležité seznámit se s klíčovými pojmy, se kterými operuje. Prvním je pragmatická kontradikce. Pragmatickou kontradikci popisuje následovně: „Tvrzení [...] jsou pragmaticky kontradiktorická, protože se pokouší udělat něco, co nemůže být provedeno, pokud jsou pravdivá.“¹¹⁹ Pragmatickou kontradikcí je tedy tvrzení, které se vyslovením stává kontradikcí. Jako příklad pragmatické kontradikce mohou sloužit dvě následující věty: „Já jsem němý“ a „Já neumím psát“. Pokud kdokoliv pronese první uvedenou větu, pak právě tím, že ji vyslovil, ji učiní kontradiktorickou, jelikož kdyby byla pravda to, co věta říká, pak by ji nikdy nemohl vyslovit. Druhá věta je příkladem pragmatické kontradikce v psaném textu.

¹¹⁷ „According to Harrison, Tarski's theory of truth must be mistaken, since (a) truth cannot be a property of sentences; (b) if truth could be a property of sentences, T-sentences would have to be necessary truths, which they are not; and (c) not only are T-sentences not necessarily true, they can be false.“ (BOISVERT, Daniel R., *The Trouble with Harrison's 'The Trouble with Tarski'*. s. 376.)

¹¹⁸ „‘things’ that are or might [...] be asserted, and which, if they were asserted, would be true regardless of the words in which they were formulated.“ (HARRISON, Jonathan. *The Trouble with Tarski. The Philosophical Quarterly*. January 1998, vol. 48, no. 190, s. 13.)

¹¹⁹ „Assertions [...] are pragmatically contradictory because they attempt to do something that could not be done if what they asserted was true.“ (Tamtéž, s. 6.)

Dalším důležitým pojmem Harrisonovy argumentace, stojícím v opozici k pragmatické kontradikci, je pragmatická nutnost. Pragmatickou nutností chápe Harrison takové tvrzení, jehož pravdivost zaručuje právě akt jeho tvrzení. Tedy jinými slovy pragmatická nutnost je takové tvrzení, které je pravdivé právě proto, že ho tvrdíme. Jako příklad pragmatické nutnosti lze použít „Henriettino jméno je ‚Henrietta‘“,¹²⁰ nebo „Já mluvím“. Podobně jako v případě pragmatické kontradikce, kdy tvrzením věty „Já jsem němý“ tuto větu právě aktem jejího tvrzení činíme kontradikcí, pragmatickou nutnost činí akt jejího tvrzení pravdivou, jako v případě, kdy vyslovím větu „Já mluvím“. V souvislosti s tímto Harrison upozorňuje, že je zapotřebí rozlišovat dva druhy vět — věty tvrdící nutnou pravdu a věty tvrdící nutně pravdu. Věty tvrdící nutnou pravdu jsou takové věty, které tvrdí něco, co nemůže být za žádných okolností nepravda. Takové věty tvrdí něco, co je logicky nutné. Oproti tomu věty tvrdící nutně pravdu jsou také vždy pravdivé, ale nikoliv proto, že to, co tvrdí je logicky nutně pravdivé. Tyto věty jsou vždy pravdivé díky způsobu, jakým jsou formulovány. „Metrová tyč v Paříži je metr dlouhá“¹²¹ je příkladem věty tvrdící nutně pravdu. To, co tato věta tvrdí není logicky nutná pravda, nicméně tato věta je formulována tak, že nemůže nastat případ, aby byla nepravdivá. Popřením této věty by došlo ke vzniku pragmatické kontradikce. Vzhledem k právě uvedené distinkci je podle Harrisona možné rozlišit dva způsoby, jak by se mohlo specifikovat T-schéma:

- a) ‚s‘ vyjadřuje nutnou pravdu tehdy a jen tehdy, když s.
- b) ‚s‘ vyjadřuje nutně pravdu tehdy a jen tehdy, když s.¹²²

Harrison je názoru, že T-schéma by mělo být interpretováno v souladu s bodem (b), jelikož význam vět jako „sníh je bílý“ se učíme aposteriorně. To by znamenalo, že T-věty nejsou nutně pravdivé věty, ale jsou to věty tvrdící nutně pravdu, a jako takové jsou pragmaticky nutné. Popřením T-věty bychom se tedy dopouštěli pragmatické kontradikce. To však znamená, že pravdivost T-vět je výsledkem toho, jak jsou formulovány a nikoliv toho, co tvrdí. „Mám dojem, že to, že ‚Sníh je bílý‘, je pravdivé tehdy a jen tehdy, když sníh je bílý, je pragmatická nutnost, vám říká velmi málo o významu ‚pravdivý‘.“¹²³ Zastánci T-schématu však zastávají (podle Harrisona) názor, že T-věty jsou nutnými

¹²⁰ „Henrietta’s name is ‘Henrietta’.“ (Tamtéž, s. 10.)

¹²¹ „The metre rod in Paris is a metre long.“ (Tamtéž.)

¹²² Srovnej tamtéž, s. 16.

¹²³ „I suspect that that ‘Snow is white’ is true if and only if snow is white is pragmatically necessary tells you very little about the meaning of ‘true’.“ (Tamtéž, s. 10.)

pravdami, což je v rozporu s Harrisonovými výsledky. Proto je třeba odmítnout T-schéma jako neadekvátní.

Kromě toho, že T-věty nejsou nutné pravdy, mohou být podle Harrisona i nepravdivé. K tomu, aby byla T-věta pravdivá, musí být obě její strany pravdivé nebo obě nepravdivé. Podle Harrisona je ale možné, aby byla pravdivá jen jedna strana, zatímco druhá by byla nepravdivá. Kdybychom mluvili o větě jazyka J^X , kde by slovo „bílý“ znamenalo černý, pak by věta

(A) Věta „sníh je bílý“ je pravdivá v J^X tehdy a jen tehdy, když sníh je bílý.

byla pragmatickou kontradikcí. Jestliže však podle kritéria materiální adekvátnosti musí mít každá adekvátní definice pravdy za následek všechny T-věty, pak vzniká problém. T-věty musí být pravdivé, což ale (A) není, proto musí být Tarského koncepce chybná.

Závěrem Harrisonova zkoumání tak je, že predikát „pravdivý“ nemůže být aplikovatelný na věty. Tudíž je nepřipustné používat predikát „pravdivý v L“, který používají zastánci T-schématu, ale místo něj by se měl používat predikát „vyjadřuje pravdu v L“. Je zde však nutné položit si otázku, zda je Harrisonova kritika adekvátní vzhledem k Tarského koncepci. Asi největším problémem Harrisonovy kritiky je, že nepoužívá verzi T-schématu, která by odpovídala konvenci T. Zatímco Tarski požaduje, aby na levé straně T-schématu bylo jméno věty a na pravé straně její překlad do metajazyka, Harrison ve své argumentaci adekvátní překlady do metajazyka nepoužívá. Příkladem budiž věta (A), na jejíž pravé straně se nevyskytuje adekvátní překlad věty z její levé strany. Za předpokladu, že metajazykem je soudobá čeština a objektovým jazykem je J^X , by věta (A) měla správně vypadat takto:

(B) Věta „sníh je bílý“ je pravdivá v J^X tehdy a jen tehdy, když sníh je černý.

Zatímco věta (B) obsahuje na pravé straně adekvátní metajazykový překlad věty na své levé straně, u věty (A) tomu tak není, tudíž věta (A) není ekvivalencí tvaru (T). Důsledkem toho je, že ani důsledky, které si z možnosti její formulace odvodil Harrison, nejsou platné. Je tedy diskutabilní, zdali je Harrisonova kritika opravdu kritikou Tarského koncepce, nebo se spíše jedná o kritiku Harrisonova vlastního konstruktů. Navíc, jak jsem již uvedl na konci kapitoly 1.2, Tarski netrval na tom, že pouze věty jsou jedinými možnými pravděnci, což opět devalvuje Harrisonovu kritiku.

10 Saul Kripke — proti bivalenci a relativizaci

Již v kapitole věnované problematice autoreferenčních vět jsem okrajově narazil na princip bivalence. V této kapitole se budu zabývat Kripkeho teorií, ve které hraje odmítnutí principu bivalence klíčovou úlohu. Cílem této kapitoly je stručně definovat Kripkeho přístup k problematice pravdy, přičemž půjde i o jisté porovnání Kripkeho a Tarského přístupů.

Stejně jako Tarski, i Kripke tvrdí, že nejvhodnější je aplikovat pravdivost na věty. Velký rozdíl je však v přístupu obou k požadavkům kladeným na adekvátní teorii pravdy. Tarski byl zastáncem bivalence, což znamená, že všechny věty mají buď pravdivostní hodnotu „pravda“, nebo „nepravda“. Nemůže se tak stát, aby měla některá věta obě pravdivostní hodnoty, nebo aby jí naopak nebylo možné pravdivostní hodnotu připsat. Ostatně právě proto, že u paradox plodících vět nejde určit jejich pravdivostní hodnota, Tarski formuloval požadavky, kterým musí jazyk, pro který se má definice pravdy formulovat, vyhovovat. Především, že tento jazyk nesmí být sémanticky uzavřený. Kripke se proti tomuto staví do opozice, když uvádí jako svůj požadavek, že „adekvátní teorie musí připouštět, aby byla naše tvrzení obsahující pojem pravdy *riskantní*“.¹²⁴ Na rozdíl od Tarského, pro kterého bylo přijatelné používat pouze klasickou logiku, Kripke hodlá pro svoje účely používat trojhodnotovou logiku¹²⁵ vyvinutou Stephenem C. Kleenem, přičemž pravdivostní predikát v ní používaný je parciální. V takovéto logice je kromě hodnot „pravda“ a „nepravda“ přípustná ještě třetí hodnota — „nedefinováno“. Aby mohl vytvořit kompletní tabulku pravdivostních funkcí pro tuto logiku, definuje Kripke, jak by vypadal průběh pravdivostních hodnot pro disjunkci. Disjunkce je pravdivá v případě, že alespoň jeden její člen má hodnotu „pravda“, nepravdivá v případě, že oba její členy mají hodnotu „nepravda“ a nerozhodnutelná v ostatních případech (tedy pokud mají oba členy hodnotu „nedefinováno“, nebo má jeden člen hodnotu „nepravda“ a druhý „nedefinováno“).¹²⁶ Použitím trojhodnotové logiky se Kripke zbavuje nepříjemných důsledků plynoucích z používání principu bivalence, jakým je např. v minulé kapitole napadané odstranění všech autoreferenčních vět (které ale nezávisí výhradně na principu bivalence). Kripke tak, stejně jako Martin, ponechává v jazyku i autoreferenční výrazy, ale jeho způsob

¹²⁴ „an adequate theory must allow our statements involving the notion of truth to be *risky*“ (KRIPKE, Saul. Outline of a Theory of Truth. *The Journal of Philosophy*. November 1975, vol. 7, no. 19, s. 692.)

¹²⁵ Právě kvůli použití trojhodnotové logiky je však Kripkeho řešení napadnutelné zesíleným lhářem.

¹²⁶ Srovnej tamtéž, s. 700.

odhalování paradoxů a přepisování pravdivostních hodnot výrazům se od Martinova postupu dost liší.

Kripkeho metoda přepisování pravdivostních hodnot výrazům je založena na další kritice Tarského. Tarski, jak jsem již zmínil v první části této práce, odmítnul, aby zkoumaný jazyk L^k obsahoval jako svoji část pravdivostní predikát aplikovatelný na výrazy tohoto jazyka samého. Pravdivostní predikát pro L^k tak musel být až v jazyku L^{k+1} , predikát pro L^{k+1} až v L^{k+2} atd. Kripke s tímto pojetím mnoha pravdivostních predikátů nesouhlasí: „Náš jazyk zajisté obsahuje pouze jedno slovo ‚pravdivý‘, nikoliv sekvenci různých frází $[\text{pravdivý}_n]$, aplikovatelných na jazyky vyšších a vyšších úrovní.“¹²⁷ Pokud by takto fungoval náš jazyk, pak by každý, kdo by používal termín „pravdivý“ k němu vlastně vždycky připojovat index (i když pouze implicitně), který by odkazoval k úrovni jazyka, ke které se daný predikát vztahuje. Navíc by pro mluvčího bylo v některých případech složité nebo dokonce nemožné zjistit, jaké úrovně by měl být jím použitý pravdivostní predikát. Ke zdůvodnění tohoto používá Kripke vlastní variantu heteroreferenčního kvantifikačního paradoxu lháře, což jsem využil výše. Na základě těchto námitek Kripke zavrhuje Tarského hierarchii metajazykových pravdivostních predikátů. Kripke konstatuje, že moderní přístupy (konkrétně Martinův a van Fraassenův) se shodují na následujících požadavcích, kterým by měla teorie pravdy vyhovovat: „měl by zde být jen jediný pravdivostní predikát aplikovatelný na věty obsahující tento predikát samotný; paradoxu by se mělo uniknout povolením pravdivostních mezer a prohlášením, že zejména paradoxní věty trpí těmito mezerami.“¹²⁸

Kripkeho cílem je tedy vytvořit koncepci, v níž bude jen jeden pravdivostní predikát aplikovatelný i na výrazy, které ho obsahují. Tento predikát není, jako v Tarského koncepci, totální, ale parciální (není tedy nutné, aby byla vždy přiřazena pravdivostní hodnota). Navíc, zatímco Tarski má na každé úrovni jazyků nový pravdivostní predikát, Kripke používá pouze jediný predikát, jehož extenze (třída pravdivých vět) i antiextenze (třída vět, které nejsou pravdivé — tj. nepravdivé věty a výrazy, které nejsou větami daného jazyka) rostou tak dlouho, dokud není dosažen tzv. *fixpoint*, neboli pevný bod. Kripkeho pravdivostní predikát je stejný pro všechny jazyky, ale v různých úrovních jazyka se mění jeho interpretace, která je dána jeho extenzí a antiextenzí. Při vzestupu do

¹²⁷ „Surely our language contains just one word ‘true’, not a sequence of distinct phrases $[\text{true}_n]$, applying to sentences of higher and higher levels.“ (Tamtéž, s. 695.)

¹²⁸ „there is to be only one truth predicate, applicable to sentences containing the predicate itself ; but paradox is to be avoided by allowing truth-value gaps and by declaring, that paradoxical sentences in particular suffer from such a gap.“ (Tamtéž, s. 698.)

vyšší úrovni jazyka dochází k rozšiřování interpretace pravdivostního predikátu — tedy na čím vyšší úrovni jazyka se nacházíme, tím větší je množství vět, které mají stanovenou pravdivostní hodnotu. Tento proces však netrvá donekonečna, ale v určitém okamžiku se dospěje do stádia, kdy se interpretace pravdivostního predikátu vzestupem do vyšší úrovni jazyka nezmění (nerozšíří) — v takovémto případě bylo dosaženo (minimálního) pevného bodu. Vzestup nad pevný bod už neobohacuje třídy vět, u nichž je určena pravdivostní hodnota, jelikož byly přiděleny pravdivostní hodnoty všem větám, kterým být přiděleny mohly. Důležitým rysem tohoto procesu je, že pokud je stanovena pravdivostní hodnota nějaké věty, pak ji má tato věta již napořád. Dalším postupem se tedy pravdivostní hodnota věty ani nezmění, ani se daná věta nemůže stát nedefinovanou. Zde by stálo za zmínku, že zatímco Tarski definoval pravdivost za pomoci pojmu splňování, Kripke dělá přesný opak: „místo abychom řekli, že $A(x)$ je splňováno objektem a řekneme, že $A(x)$ se stává pravdivým, když je proměnná nahrazena jménem a .“¹²⁹

Věty, které mají pravdivostní hodnotu v minimálním pevném bodu Kripke označuje jako uzemněné, jestliže věta nemá pravdivostní hodnotu v pevném bodu, pak je neuzemněná. To, jestli je věta uzemněná nebo nikoliv je mnohdy záležitostí empirických faktů.¹³⁰ Pokud je daná věta výpovědí o jiných větách, pak její pravdivost (a uzemněnost) závisí na pravdivosti právě těchto vět. Jako příklad může posloužit věta Buridanova paradoxu pronesená Sókratem „Vše, co řekl Platón, je nepravdivé“. Abychom mohli určit, zda je tato věta uzemněná nebo nikoliv (tedy jestli má v pevném bodu pravdivostní hodnotu), pak bychom museli zjistit, jaké věty Platón řekl a jaké jsou jejich pravdivostní hodnoty. Sókratova věta se tak může stát uzemněnou pouze tehdy, když budou uzemněné všechny věty pronesené Platónem. Rysem paradox plodících vět je, že takové věty nemají pravdivostní hodnotu v žádném pevném bodu.

Toto byla velmi stručná charakterizace Kripkeho teorie pravdy. Její přesnější popis lze najít např. u Scotta Soamese.¹³¹ Závěrem však zbývá dodat několik poznámek a shrnutí. Jak jsem uvedl na začátku kapitoly, Kripke napadá dva fundamenty Tarského koncepce. Prvním je princip bivalence — Kripkeho odpověď spočívá ve využití trojhodnotové logiky. I když tím uniká některým nepříjemným důsledkům Tarského teorie, zároveň se tím vystavuje napadení zesíleným lhářem. Druhou napadenou součástí Tarského koncepce

¹²⁹ „instead of talking of $A(x)$ being satisfied by an object a , we can talk of $A(x)$ becoming true when the variable is replaced by a name of a .“ (Tamtéž, s. 705.)

¹³⁰ Tuto koncepci rozvíjí, spolu s celou tzv. Kripke-Fefermanovou teorií, Tim Maudlin v MAUDLIN, Tim. *Truth and Paradox: Solving the Riddles*. Oxford: Clarendon Press, 2006.

¹³¹ Viz SOAMES, Scott. *Understanding Truth*. s. 181–195.

je hierarchie pravdivostních predikátů. Zatímco Tarski má pravdivostní predikát pro každý jazyk v jazyku o řád vyšším, Kripkeho řešení spočívá ve snaze vytvořit jediný parciální pravdivostní predikát, který je navíc aplikovatelný i na sebe sama. Kripke k tomuto dodává: „*všechny pravdivostní predikáty Tarského konečné hierarchie jsou definovatelné uvnitř \mathcal{L}_σ a všechny jazyky té hierarchie jsou podjazyky \mathcal{L}_σ* “¹³² Ačkoliv Kripke zredukoval počet pravdivostních predikátů na jeden, který se nenachází v metajazyku, stejně se mu nepodařilo úplně se vyhnout všem důsledkům Tarského teorie, jelikož pojmy jako „uzemněnost“ a „paradoxní“ se nacházejí v metajazyku.¹³³

¹³² „*all the truth predicates of the finite Tarski hierarchy are definable within \mathcal{L}_σ , and all the languages of that hierarchy are sublanguages of \mathcal{L}_σ* “ (KRIPKE, Saul. *Outline of a Theory of Truth*. s. 710.)

¹³³ Srovnej FIELD, Hartry H. *Saving Truth from Paradox*. Oxford: Oxford University, 2008. s. 222–223.

Závěr

Záměrem první části mé práce bylo podat stručný výklad Tarského sémantické koncepce pravdy, tak, jak je prezentována v dvou jeho nejznámějších statích [CTFL] a [SCT]. Hlavním zdrojem byla sice první jmenovaná stať, ale jelikož byla v některých ohledech Tarským možná až příliš zjednodušená, pokusil jsem se svůj výklad doplnit za pomoci informací ze statí druhé. To se týká především kapitol 2.1, 2.3 a 4. Oproti tomu kapitola 2.2 byla prakticky celá inspirována sekundární literaturou. Snažil jsem se podat přehled základních principů Tarského koncepce tak, aby byly zmíněny všechny její důležité aspekty, včetně těch, které nejsou v [SCT] explicitně uvedeny, přičemž mým cílem byla snaha, aby tento výklad byl co možná nejsrozumitelnější.

Stejně jako Tarski v [SCT], i já se rozhodl věnovat druhou část své práce kritice Tarského sémantické koncepce pravdy. Mým záměrem nebylo podat úplný výčet kritik vůči této koncepci vznesených, jelikož takový výčet by sám vydal na velmi obsáhlou práci. Místo toho jsem se pokusil vyložit alespoň ty z nich, které napadají nosné myšlenky Tarského pojetí, a zároveň jsem se u některých pokusil poukázat na jejich nedostatky.

První dvě kritiky byly zaměřeny proti hierarchiím metajazyků a daly by se proto vztáhnout nejenom na Tarského koncepci. V případě těchto námitek jsem k poukázání na jejich nedostatky využil poznatky dr. Jiřího Raclavského.

Druhá kritika byla zaměřena proti Tarského ukvapenosti v odmítání autoreferenčních vět. Tarski je zamítá na základě principu bivalence a kritérií, kterým nesmí vyhovovat žádný jazyk, pro který má být vytvořena adekvátní definice pravdy. Tuto kritiku přednesl spolu s návrhem na řešení problému autoreference R. L. Martin.

Kritika z pera Hanse Herzbergera je namířena proti Tarského odsouzení hovorových jazyků coby nekonzistentních. Herzberger vychází z přesvědčení, že jazyk může být nekonzistentní pouze tehdy, pokud je nekonzistentní množina analytických vět v daném jazyku obsažených. Proti Tarskému staví *reductio ad absurdum*, jehož cílem je dokázat, že žádný jazyk nemůže obsahovat nekonzistentní množinu analytických vět a být tak nekonzistentní.

Tarského rozhodnutí aplikovat termín „pravdivý“ na věty kritizuje Jonathan Harrison. Ve své kritice napadá Tarského T-schéma a snaží se dokázat, že ekvivalence tvaru (T) mohou být pravdivé, což by mělo za následek vyvrácení Tarského koncepce. Harrison se ale ve své kritice, podle Daniela Boisverta, dopouští dezinterpretace Tarského a nepředstavuje tak pro jeho koncepci nebezpečí.

Saul Kripke odmítá princip bivalence a místo pravdivostních predikátů pro každou úroveň jazyků zavádí jediný. Použitím trojhodnotové logiky se však vystavuje zesílenému lháři. Navíc ani jediný parciální pravdivostní predikát ho „nezachrání“ před nutností používat metajazyky.

I když Tarského koncepce rozhodně není bezchybná, jak ukázaly i tyto kritiky, přesto je i nadále velmi inspirativní a nezbyvá mi než spolu s Kripkem dodat: „Duch Tarského hierarchie je stále s námi.“¹³⁴

¹³⁴ „The ghost of the Tarski hierarchy is still with us.“ (KRIPKE, Saul. Outline of a Theory of Truth. 714.)

Použitá a citovaná literatura

- 1) ARISTOTELÉS. *Metafyzika*. Přeložil Antonín Kříž. Praha: Nakladatelství Petr Rezek, 2003.
- 2) BOISVERT, Daniel R. The Trouble with Harrison's 'The Trouble with Tarski'. *The Philosophical Quarterly*. July 1999, vol. 49, no. 196, s. 376–383.
- 3) CARNAP, Rudolf. *Problémy jazyka vědy: (výbor prací Rudolfa Carnapa)*. Praha: Svoboda, 1968.
- 4) DAVIDSON, Donald. The Folly of Trying to Define Truth. *The Journal of Philosophy*. June 1996, vol. 93, no. 6, s. 263–278.
- 5) FIELD, Hartry H. *Saving Truth from Paradox*. Oxford: Oxford University, 2008.
- 6) GÓMEZ-TORRENTE, Mario. The indefinability of truth in the "Wahrheitsbegriff". *Annals of Pure and Applied Logic* 126. 2004, s. 27–37.
- 7) HARRISON, Jonathan. The Trouble with Tarski. *The Philosophical Quarterly*. January 1998, vol. 48, no. 190, s. 1–22.
- 8) HECK, Richard G. Tarski, Truth, and Semantics. *The Philosophical Review*. October 1997, vol. 106, no. 4, s. 533–554.
- 9) HERZBERGER, Hans G. The Truth-Conditional Consistency of Natural Languages. *The Journal of Philosophy*. February 1967, vol. 64, no. 2, s. 29–35.
- 10) KIRKHAM, Richard L. *Theories of Truth: A Critical Introduction*. Cambridge: MIT Press, 1995.
- 11) KRIPKE, Saul. Outline of a Theory of Truth. *The Journal of Philosophy*. November 1975, vol. 7, no. 19, s. 690–716.
- 12) LEVISON, A. B. Logic, Language, and Consistency in Tarski's Theory of Truth. *Philosophy and Phenomenological Research*. March 1965, vol. 25, no. 3, s. 384–392.
- 13) LYNCH, Michael B. *The Nature of Truth: Classic and Contemporary Perspectives*. MIT Press: Cambridge, 2001.
- 14) MARTIN, Robert L. Toward a Solution to the Liar Paradox. *The Philosophical Review*. July 1967, vol. 76, no. 3, s. 279–311.
- 15) MAUDLIN, Tim. *Truth and paradox: Solving the Riddles*. Oxford: Clarendon Press, 2006.
- 16) MILNE, Peter. Tarski on Truth and Its Definition [online]. 1997 [cit. 2010-02-12]. Dostupné z: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.103.712&rep=rep1&type=pdf>.
- 17) RACLAVSKÝ, Jiří. Lhářský paradox, význam a pravdivost [online]. 2009 [cit. 2010-02-12]. Dostupné z: http://www.phil.muni.cz/~raclavsk/texty/lharske_paradoxy.pdf.
- 18) RACLAVSKÝ, Jiří. Paradox lháře a jeho řešení [online]. 2007 [cit. 2010-02-12]. Dostupné z: http://www.phil.muni.cz/~raclavsk/texty/paradox_lhare.pdf.
- 19) SINISI, Vito F. Tarski on the Inconsistency of Colloquial Language. *Philosophy and Phenomenological Research*. June 1967, vol. 27, no. 4, s. 537–541.

- 20) SOAMES, Scott. *Understanding Truth*. Oxford, New York: Oxford University Press, 1999.
- 21) SUPPES, Patrick. Philosophical Implications of Tarski's Work. *The Journal of Symbolic Logic*. March 1988, vol. 53, no. 1, s. 80–91.
- 22) TARSKI, Alfred. Sémantická koncepce pravdy a základy sémantiky. In *Logika 20. století: Mezi filosofií a matematikou: výbor textů k moderní logice*. Praha : Filosofia, 2006. s. 135–176.
- 23) TARSKI, Alfred. The Concept of Truth in Formalized Languages. In *Logic, Semantics, Metamathematics*. Oxford; New York: Oxford University Press, 1956. s. 152–278.
- 24) TARSKI, Alfred. The Semantic Conception of Truth: and the Foundations of Semantics. *Philosophy and Phenomenological Research*. March 1944, vol. 4, no. 3, s. 341–376.